

LOCKE

Epistemology & Ontology

Michael Ayers

An Essay concerning (3)

*The Understanding, Knowledge,
Opinion
Belief & Assent*

1 § Since it is the Understanding that ~~sets~~ ^{gives} man ~~about~~ ^{with} the sensible beings & gives him all that Dominion wh^{ch} he hath over them it is certainly a subject even for its ~~notion~~ ^{notion} w^{ch} may labour to enquire into, & wh^{ch} perhaps hath been ~~less~~ ^{less} seriously considered than yr way of the thing & yr manner of it hath been seem^{ing} to require. The understanding like the eye wh^{ch} it makes use of see^s & perceiv^{ing} all other things, hath no notice of its self, & it must require art & pains to set it at a distance from its self & make it its own object. But what was the difficulty that lay in yr way of this ~~enquiry~~ ^{enquiry}. What was it but a heap of

ROUTLEDGE

THE ARGUMENTS OF THE PHILOSOPHERS

Edited by Ted Honderich

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

LOCKE

John Locke is the greatest English philosopher. *An Essay Concerning Human Understanding*, one of the most influential books in the history of thought, is his greatest work. In this study the historical meaning and philosophical significance of Locke's *Essay* are investigated more comprehensively than ever before.

Locke is the result of many years' work. Already influential through his writings on the history of philosophy, Michael Ayers has produced a study of Locke's *Essay* which is an important contribution to both scholarship and philosophy. This study was originally published in two volumes, *Epistemology* and *Ontology*. This paperback edition has within its covers the full text of both volumes.

Michael Ayers is a Fellow of Wadham College, Oxford, and a Lecturer in Philosophy at Oxford University.

The Arguments of the Philosophers

EDITOR: TED HONDERICH

*Grote Professor of the Philosophy of Mind and Logic,
University College, London*

The purpose of this series is to provide a contemporary assessment and history of the entire course of philosophical thought. Each book constitutes a detailed, critical introduction to the work of a philosopher of major influence and significance.

Already published in the series:

**Augustine* Christopher Kirwan

**J.L.Austin* Geoffrey Warnock

Ayer John Foster

**Bentham* Ross Harrison

Bergson A.R.Lacey

**Berkeley* George Pitcher

Butler Terence Penelhum

**Descartes* Margaret Dauler Wilson

**Dewey* J.E.Tiles

†*Gottlob Frege* Hans Sluga

Hegel M.J.Inwood

Hobbes Tom Sorell

Hume Barry Stroud

Husserl David Bell

William James Graham Bird

**Kant* Ralph C.S.Walker

Kierkegaard Alastair Hannay

**Karl Marx* Allen Wood

†*Meinong* Reinhart Grossman

**John Stuart Mill* John Skorupski

**G.E.Moore* Tom Baldwin

- *Nietzsche* Richard Schacht
Peirce Christopher Hookway
**Plato* Justin Gosling
**Karl Popper* Anthony O'Hear
**The Presocratic Philosophers* Jonathan Barnes
**Thomas Reid* Keith Lehrer
**Russell* R.M.Sainsbury
†Santayana Timothy L.S.Sprigge
**Sartre* Peter Caws
**Schopenhauer* D.W.Hamlyn
**Socrates* Gerasimos Xenophon Santas
Spinoza R.J.Delahunty
**Wittgenstein* Robert J.Fogelin

**available in paperback*

†no longer available

LOCKE

Epistemology and Ontology

Michael Ayers



London and New York

First published as two volumes 1991 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

First published in paperback 1993

This edition published in the Taylor & Francis e-Library, 2005.

“To purchase your own copy of this or any of Taylor & Francis or Routledge’s collection
of thousands of eBooks please go to www.eBookstore.tandf.co.uk.”

Routledge is an International Thomson Publishing company

© 1991 Michael Ayers

All rights reserved. No part of this book may be reprinted or reproduced or utilised 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 Cataloguing in Publication Data

Ayers, Michael, 1935—Locke/Michael Ayers. p. cm.—(Arguments of the philosophers)
Two previously hard bound vols. are combined in 1. pbk. vol. Includes bibliographical
references and index

ISBN 0-203-00641-0 Master e-book ISBN

ISBN 0-203-17431-3 (Adobe eReader Format)

ISBN 0-415-10030-5 (Print Edition)

1. Locke, John, 1632–1704. Essay concerning human understanding. 2. Knowledge,
Theory of-History-17th century. 3. Ontology-History-17th century. I. Title. II. Series.

B1294.A94 1993

192–dc20

93–15180

CIP

Contents

VOLUME I EPISTEMOLOGY

Acknowledgements	x
Introduction	1
PART I IDEAS	
1 Introduction to Part I	11
2 Ideas and compositionism in traditional logic	16
3 Ideas and epistemology before Locke	22
4 Simple and complex ideas	31
5 Ideas as images	38
6 Ideas as intentional acts and ideas as intentional objects	45
7 Ideas as natural signs	52
8 Ideas: a summary	59
9 Reflections on the structure of thought	62
PART II KNOWLEDGE AND BELIEF	
10 Introduction to Part II	70
11 The degrees of knowledge and the role of method	76
12 Other divisions of knowledge	83
13 Probability and the nature of 'assent'	90
14 The grounds of probability	98
15 Reflections on the definition of knowledge	109
16 Belief and rationality	127

PART III PERCEPTUAL KNOWLEDGE

17	Introduction to Part III	133
18	The authority and limits of ‘sensitive knowledge’	135
19	Does perceptual knowledge have independent authority?	145
20	Does perceptual knowledge have a firm boundary?	151
21	The scope of perceptual knowledge	158
22	Two modern approaches to sensation	170
23	Private language and secondary qualities	183

PART IV PARTICULARS, UNIVERSALS AND INTUITIVE KNOWLEDGE

24	Introduction to Part IV	194
25	Locke’s arguments on space and time in context	196
26	Thought about particulars	209
27	Locke’s theory of universal knowledge in context	213
28	Abstraction and the ideal of precision	228
29	Intuition and innate knowledge	232
30	Locke on meaning and some modern criticisms	237
31	Reflections on understanding and imagination	244
32	Necessity, reason and language	255
33	Conclusions of Volume I	266
	Notes	270
	Bibliography	297
	Index	301

VOLUME II ONTOLOGY

Introduction	1
--------------	---

PART I SUBSTANCE AND MODE

1	Introduction to Part I	15
2	Substance, essence and accidents before Locke	18
3	Our complex ideas of substances and the idea of substance in general	31
4	Substance and real essence, matter and spirit, and the obscurity and confusion of the idea of substance	39
5	Substance, mode and the argument from language	51
6	Species and their names in the corpuscularian world	65
7	Are there real species?	78
8	Locke on the difference between substances and modes	91
9	Reflections on the notion of substance	110
PART II GOD, NATURE AND THE LAW OF NATURE		
10	Introduction to Part II	131
11	Forms of mechanism before Locke	135
12	The form of Locke's mechanism	142
13	Reflections on rationalism, empiricism and mechanism	154
14	The existence of God	169
15	The Law of Nature and human freedom	184
16	Reflections on Locke's ethics	196
PART III IDENTITY		
17	Introduction to Part III	205
18	Locke on 'masses of matter'	207
19	Locke on living things	216
20	Forms of material unity	229
21	Artificial and other problematic objects	239
22	Personal identity before the <i>Essay</i>	254
23	Locke's theory of personal identity	260
24	Contemporary reactions to Locke's theory	269
25	Neo-Lockean and anti-Lockean theories of personal identity in analytic philosophy	278

26	Conclusions of Volume II	293
	Notes	297
	Bibliography	329
	Index	333

Acknowledgements

I am grateful to those who gave the appropriate permissions to incorporate, as follows, some variously modified and amended parts of previously published articles in the present work. In Volume I: 'Locke's Logical Atomism', *Proceedings of the British Academy*, Volume LXVII, 1981 (as parts of chapters 1, 2 and 4), by permission of the British Academy; 'Are Locke's "Ideas" Images, Intentional Objects or Natural Signs?', *Locke Newsletter* 17, 1986 (as parts of chapters 5, 6, and 7), by permission of the Editor, Roland Hall; 'Locke's Doctrine of Abstraction: Some Aspects of its Historical and Philosophical Significance', from *John Locke: Symposium Wolfenbuttel 1979*, edited by R. Brandt (Walter de Gruyter, Berlin 1981) (as the main part of chapters 27 and 28); 'Some Thoughts', *Proceedings of the Aristotelian Society* 1972 (as part of chapter 31), by courtesy of the Editor; 'Reason and Psycholinguistics', from *Wisdom: Twelve Essays*, edited by R. Bambrough (Basil Blackwell, Oxford 1974) (as chapter 32). In Volume II: 'Locke versus Aristotle on Natural Kinds', *The Journal of Philosophy* 1981 (as the main part of chapter 6); 'Substance: Prolegomena to a Realist Theory of Identity', *Journal of Philosophy* 1991 (as part of chapter 8, and the main part of chapter 9); 'Mechanism, Superaddition and the Proof of God's Existence in Locke's *Essay*', *The Philosophical Review* 90, 1981 (as the main part of chapter 14), by permission of the publisher.

I have used Peter Nidditch's Clarendon edition of *An Essay concerning Human Understanding*, keeping the seventeenth-century spelling and punctuation which may irritate some as much, I suppose, as it may charm others. I have done so neither to irritate nor charm, but in order to remain faithful to a distinguished piece of scholarship. In other cases too I have followed the edition cited. I have everywhere used the excellent translation of Descartes' *Philosophical Writings* by John Cottingham, Robert Stoothoff and Dugald Murdoch with only this present acknowledgement, since they sufficiently indicate the cited pagination of the edition of Adam and Tannery. Translations of other authors are as indicated or are my own. The form of references is explained at the beginning of the footnotes.

I have many more personal obligations. It is not easy to identify one's deepest intellectual debts to the living, and the effort to do so now is instructive. Given

how I see philosophy, philosophical argument and philosophical understanding, it seems clear that my first acknowledgement should be the same now as twenty-five years ago, to my graduate supervisor John Wisdom. For I have to acknowledge that, through what have seemed at the time quite radical shifts in interest, method and belief, my thinking about philosophy and its history has kept within a general view of the clash of theories and of the ordinary intermingling of philosophical insight with error which has much in common with Wisdom's. Some of my own arguments also now remind me, not only of Wisdom's dissatisfaction with Wittgenstein's tendency to view traditional philosophy as mere illusion, but of his general opposition to what P.F. Strawson has described as the reductive rage. This relationship may be evident, beneath admittedly large differences of approach and style, both when occasional reference is made to Wisdom's writings and elsewhere, as in the structure employed in the two arguments which are central to Volume I and Volume II respectively: the general explanations of the concepts of knowledge and of substance. The form of those explanations occurred to me only relatively late in the day, but the seed was surely sown, very many years ago, by John Wisdom. I hope that the differences are not such that he would entirely disown the plant.

Also on the side of philosophy I must mention (without having any idea of what he would think of the present work) Strawson himself, whose books did much to persuade me that answers to the really important questions of metaphysics are most likely to be found in an adequate reassessment of traditional theory of substance. David Wiggins, despite our differences of view, is another who has influenced my conception of these questions, and I am grateful for his active and friendly interest and encouragement. A class given with him and Michael Woods provided an essential opportunity to subject my ideas on identity to expert and principled criticism without which they would have been, I believe, less clear and less consistent. Like everybody else with realist intuitions, I have learnt a lot from various writings of Saul Kripke and Hilary Putnam. And, like everybody else, in my thinking about knowledge I owe much to the discussion sparked off by the Gettier Problem (less only than to traditional epistemology). Other specific debts are recorded in the text.

This is a work which is about philosophy through being a historical and critical analysis of a particular philosophy constructed and argued at a particular time and place. My debts are therefore no less deep to those historians of philosophy who have influenced and encouraged my study of the seventeenth century. When preparation for lectures on the Empiricists long ago made me dissatisfied with the (as it seemed to me) grossly unhistorical approach to texts which was more or less universal among analytic philosophers in the sixties, the first informed encouragement I received was given with great vigour and generosity by Rom Harré, perhaps the only member of the Oxford Sub-faculty of Philosophy at the time who was qualified to give it. I sought him out after coming across his book *Matter and Method*, and he soon had me reading Robert Boyle. He also introduced me to John Yolton, who was by luck visiting Oxford. John's

remarkable knowledge of the primary philosophical literature and his infectious enthusiasm carried me into the then small world of serious Locke studies, a world which already owed much to him and has come to owe much more. I began to give classes on the history and historiography of philosophy, in that first year with John Yolton and then on my own, and some who attended them have continued to give support and friendship over the years.

Since then discussions or exchanges with many others have helped me to develop ideas about the history of philosophy and knowledge of Locke in particular. They include Roger Ariew, Reinhard Brandt, James Buickerood, Alan Gabbey, Daniel Garber, Ian Harris, Gary Hatfield, Timothy Hodgson, Edmund Leites, Thomas Lennon, Menno Lievers, Edwin McCann, John Milton, Gabriel Nuchelmans, David Owen, Richard Popkin, John Rogers, Henry Schankula, Sandy Stewart, Ian Tipton, Udo Thiel, Wolfgang von Leyden, Milton Wachsberg, Roger Woolhouse and Jean Yolton. I have a special obligation to John Dunn, Jonathan Rée and Quentin Skinner, whose writings and occasional (or, in Jonathan's case, quite regular) conversation over the years have supplied a necessary stimulus to my thinking about method. Certain meetings at which I had the opportunity to expound some of my ideas and to hear the ideas of others have also been enormously helpful to me. I would like to thank those, too many to mention by name, who gave such help, as organizers or participants, at the Locke Workshop at Edmonton in 1976 organized by Peter Schouls and John Yolton, the Wolfenbuttel Symposium on Locke in 1979 organized by Reinhard Brandt, the Locke Workshop at Rutgers in 1983 organized by John Yolton, the conference on the history and historiography of philosophy at Lancaster University in 1983 organized by Stuart Brown and Alan Holland and sponsored by the Royal Institute of Philosophy, the one-day conference on Locke at Rhodes House, Oxford, in 1987, organized by John Milton on behalf of the British Society for the History of Philosophy, and the outstanding six-week Summer Institute on Early Modern Philosophy at Brown University in 1988, organized by Daniel Garber on behalf of the Council for Philosophical Studies with very generous support from the National Endowment for the Humanities.

In addition, John Ackrill, Ian Crombie, Quassim Cassam, Colin McGinn and John Milton have all read parts (in John Milton's case, the whole) of the work in draft and have given helpful comment and advice, bibliographical, historical or philosophical, which have led to the removal of some of its imperfections. More were removed as a result of the careful and sensitive queries and suggestions of the copy-editor for Routledge, Antonia Pledger. Those many that remain (as is always said on such occasions, but as certainly needs to be said on this one) are due only to the author. Over the years I have had secretarial assistance from a variety of sources, but several people deserve particular thanks. From the Sub-faculty of Philosophy, Dorothy Cuninghame, Jo Bickerstaffe, Anne Shackle and our supremely helpful Administrator, Jane Hardie. From my College, Maureen Clark and Margaret Matthews have both been especially helpful, the latter

showing great and good-humoured patience in the awkward preparation of the final drafts from variously word-processed material.

Finally, my work owes more to my wife, Delia Ayers, and to my son and daughter, Jason and Selina Ayers, than perhaps they have been aware. Without them life would have lacked its bloom. To them this book is dedicated with love.

Volume I

Epistemology

Introduction

Perhaps the most profound and difficult question in philosophy concerns the nature of philosophy itself. One way of raising it is to ask why the history of philosophy matters so much to philosophers. Almost everywhere its study is recognized as an integral, even essential part of a philosophical education, and it is normal for it to be pursued in the same university departments and by many of the same people as engage directly in the criticism and production of current philosophical theory. It is only necessary to compare philosophy in this respect with chemistry or biology or mathematics to see that the relationship between the historiography and the practice of philosophy is a peculiar one. What kind of subject can it be that has been so bound up with its own history, ever since it was old enough to have a history? Not that an explicit conception of ‘the history of philosophy’ is nearly as old as philosophy itself. Yet philosophers have for a very long time felt it appropriate to draw on the past, to align themselves with or against long-standing traditions, to revive what has previously been discarded, to engage in dialogue, sympathetic or hostile, with the long dead. Why?

Not everyone would agree that the tie between philosophy and its historiography is either inevitable or desirable. There are both philosophers and historians who are tired of the marriage and would like a divorce. Many historians see an active interest in philosophical questions as an inevitably distorting factor in the enterprise of uncovering the content of past thought and mapping its place in the past as a whole. Some philosophers regard any serious respect for the paradigms from the past as an obstacle to a properly scientific philosophy. One of the purposes of the present work, unlikely enough to be achieved, is to persuade both parties that they are wrong.

A sceptical explanation of the importance of history to philosophy is that it is only through history, and only through that kind of motivated interpretation of past writing which constitutes the construction of a tradition or canon, that the illusion can be created that philosophy is a distinct discipline or branch of knowledge. ‘Philosophy’, according to this line of argument, is naïvely taken to be that discipline in which all the famous philosophers have been engaged: it is assumed to have a unity and a firm boundary only because we construct for ourselves a self-flattering history of coherent and consequential progress towards

what we regard as the philosophical insights of the present, a history of man's increasing grasp of some set of alleged truths. Yet critical examination of any such story (it may be argued plausibly enough) and a properly historical reconstruction of the meaning of the revered texts in the contexts in which they were written, will reveal that 'philosophical' writers have had at different times and places a variety of quite disparate aims and interests, and that the presuppositions of their theorizing have undergone quite radical alteration and replacement during the history of 'philosophy'. Doesn't it follow that there *is* no one thing on which 'philosophers' have been engaged? Perhaps this fact has only been disguised by the readiness of each generation to take off from a polemical misrepresentation of their predecessors' doctrines, or to bend existing terminology to new purposes. Perhaps it continues to be disguised only because exponents of the tradition conveniently ignore the connections, motivation and presuppositions of the 'philosophical' theories of the past, which were more closely related to issues, now dead, in contemporary theology, science, politics or the like than to anything that much concerns 'philosophers' today. Or so, at least, it may seem.

Such scepticism about the very existence of philosophy often lies behind the hostility of historians to the critical examination of texts from a philosophical point of view. It is also, of course, consonant with the long fashionable philosophical theory, which has taken a variety of forms in the present century, that philosophy in the traditional sense is as dead as alchemy. I do not share that scepticism, but I have considerable sympathy with the view that philosophical commentary on philosophical texts has commonly distorted their meaning. Philosophers have indeed proved very liable to import their own interests and preconceptions into their readings of the great, and to create a mythological past, golden or dark, in justification of present ways of doing things. They have done so, in one way or another, for a very long time, and custom has bestowed a spurious respectability on the practice. Some even see it as a necessary task for each generation to rewrite the history of philosophy according to its own philosophical lights. Yet we have the same duty here as elsewhere in historiography to try to cut through the barriers set up by the immediate context of our own thinking and to reconstruct and judge the past as it was. No doubt our reconstruction will be less than perfect. We shall never be able to see things exactly as they might have been seen by this or that seventeenth-century intellectual. No doubt, too, we must start, in our very first readings, from where we are, bringing our own philosophical intuitions or grasp of the philosophical possibilities to the interpretation of the text. Yet we should remain open to the likelihood that our first impressions of its meaning will prove wrong, even wildly wrong.

The way to correct such mistakes, of course, or even to get ourselves into the position of having anything at all worth saying about a text, is to read more. We must try to reconstruct the context in which the work was written, we must equip ourselves to recognize allusions, borrowings, oppositions, party lines, and we must not restrict our reading according to a preconception of what 'philosophy'

is. Then, very often, arguments and doctrines which at first seemed merely disparate will be seen to belong to the same network of argument. Very often, too, some will appear less consonant with present-day ideas than we were at first disposed to believe. All historical inquiry necessarily selects from the indefinite quantity of evidence available about the past. Yet that does not mean that we have to decide on principles of selection blind, before inquiry has begun. We should allow ourselves to be led by the material itself. We should be prepared to pursue indications as to what is likely to cast light on meaning without caring too much about its immediate philosophical interest to ourselves, or even whether it is 'philosophy' at all. The task of uncovering the network of influences is no doubt potentially endless in theory, but in practice an end will be reached as a result of some sense of diminishing returns, or of having bitten off as much as we personally can chew. Even the latter, in so far as it is due to chance, may be better than a limitation achieved by filtering out from the text everything which seems to us to fall under 'theology' or 'science' or some other heading different from 'philosophy' as we now employ the term.

Some may draw the conclusion from all this that philosophy should after all disentangle itself from the historiography of philosophy. If it is intellectually disreputable for a philosopher to be satisfied with a direct response to a classic text, uncorrected by the sort of scholarly inquiry into its context which alone permits us to approach the meaning it had for its author, then it may seem that most of the interest that philosophers have had in philosophy's past has been intellectually disreputable. But if the alternative is for philosophers to become historians, what profit is there for them in that?

The present work, which comprises two complementary but free-standing volumes, has been written as some kind of answer to that question. Since it is intended primarily for philosophers and students of philosophy, it is probably not historical enough in its purposes to satisfy many intellectual historians. I have not attempted a comprehensive account of the development of the doctrines I discuss, or even of the development of Locke's thought, although I have made a few suggestions as to both. What I have chiefly tried to do, however, is perhaps more than enough for one work. That is, to expound the main arguments of a single text, *An Essay concerning Human Understanding*, holding them up against as much antecedent and contemporary writing, including other writings by Locke himself, as has seemed necessary to explain their meaning and point. The selection of such material from outside the text has sometimes been fairly arbitrary, and its inclusion does not constitute a claim that Locke was responding to the particular writing cited rather than to some other expression of similar ideas. Sometimes, however, I have dared to make such a claim. Very often the comparisons represent, in reverse, my own progress to what I take to be understanding. Those passages of Aristotle, Lucretius, Descartes, Digby, Gassendi, Hobbes, Boyle, the Port Royal *Logic*, Malebranche and the like by which the scene is set were often, read or reread with the text of the *Essay* specifically in mind, the very catalysts which transformed my interpretation of this

or that argument or assertion. At the same time my efforts to understand the *Essay* have been the means by which patterns in previous thought have been made apparent to me. The result is probably too historical to please many philosophers. Yet there was no possibility of my simply saying what I take to be Locke's meaning without also saying what antecedents of the *Essay* have led to my interpretation. For, in so far as the interpretation is right, Locke's meaning is defined by the relation of what he wrote to those antecedents.

Some who agree that in order to reveal meaning a text has to be placed in a context nevertheless argue that the notion of meaning supplies no ground for favouring the context contemporary with the text over a context more immediate and familiar to the modern reader. With respect to the meaning or content of any assertion, it is suggested, the speaker's response to a question which had not occurred to him at the time might be just as revealing as his response to a question which had been at the forefront of his mind. Indeed the need to respond to a question of the former type can reveal to a speaker himself the meaning of his own assertion, as he explores its implications or removes indeterminacies. As Richard Rorty has asked in a recent discussion of the historiography of philosophy, since the construction of a speaker's meaning always involves 'finding out what he would have said in reply to questions about what he said previously', why treat the text's antecedents as the privileged source of such questions? Don't we reveal Locke's 'meaning' as readily by considering what answers he might have given to our own twentieth-century questions as by considering the replies he might have given to the unanswered criticisms actually formulated by Leibniz in Locke's lifetime or soon after his death?

A short rebuttal of this sophistical argument is that meaning, although not simply a matter of what the writer intended, is very much bound up with intentions, while what was intended is similarly bound up with what was going on in the writer's head and, indeed, his consciousness at the time. And in so far as meaning is not determined by intention, it is at least largely determined by something equally time-bound; the language employed by the writer. Both these considerations give sufficient reason to favour the intellectual context of the work we are studying over our own intellectual context when it comes to giving an account of its meaning. Neither is outweighed by the truism that interpretation requires us to bring the text into relation with ourselves. Admittedly, in interpreting any assertion we need to understand its maker to be talking some kind of sense, to be saying something which, although we may not believe it to be true, at least appears to us to have some kind of intelligible foundation or motivation from what we can justifiably consider to be the asserter's point of view. But that does not, of course, mean that intelligent interpretation is possible without any attempt to reconstruct the speaker's or writer's point of view.

A related point is this: every particular hypothetical question concerns the outcome of some actual, given *situation* in supposed or postulated *circumstances*. But to suppose Locke's being faced with Quine's principle of the indeterminacy of translation or with quantum mechanics is to postulate circumstances too far

from the given situation. The supposition takes us too far from the thought of the *Essay* for there to be any determinate or persuasive answer to the question, 'What would Locke have said about these things?' To give an answer would be to play a game with no rules. That is not to say that it will be in principle impossible to find a persuasive answer to *any* such question asked from the point of view of the present, but how far it is possible will depend on the distance between the present and the past in the relevant respect. Nor does it mean that the *Essay* supplies grounds for determinate answers to all the questions that were or might have been raised as to its meaning by Locke's contemporaries, even less innovative contemporaries than Leibniz. In the *Essay*, as (one supposes) in every philosophical text in existence, there are real indeterminacies of meaning which are more than illusions created by anachronistic questions. Several such indeterminacies, for example, attach to the chapter on identity, a chapter in which, as perhaps in no other, Locke was exploring unmapped country.

A second task attempted in the present work, no doubt with imperfect success, is that of assessing Locke's arguments, not just one by one, but as a whole in the only way in which a philosophical system can be at all adequately assessed: by means of a response as systematic as the philosophy under consideration. To say that is of course to assume that the sceptic is wrong and that the notion of systematic philosophy is not a delusion. In support of this assumption there is little to be said prior to substantive argument. The proof of the pudding is in the eating. But some things can be said. For example: there is a natural world in which we ourselves exist; we have thoughts and experiences (whatever they are); we have knowledge and beliefs; we have scientific theories and mathematics; we possess language. All these things stand in some general relationships to one another, of which it must in principle be possible to give some explanatory account. It may be that language is a necessary condition of mathematics, or that every belief necessarily embodies some theory, or that mathematical truth is independent of any truth about the world. Yet if we possessed a satisfactory understanding of all such relationships we would be able to give correct and systematic answers to the central questions of epistemology and ontology. Conceivably those correct answers would not look much like any traditional answers, but that is not the point. The point is that, whatever their various motives in getting into the game, the famous 'philosophers' have treated such interconnected matters as the relations between thought, language and reality, or the relations between reason, emotion, belief and action; while at least many of the other things with which they have been concerned have been intelligibly connected to that central core. Members of the tradition responsible for the repeated announcements of the end of 'metaphysics' or 'epistemology' have been no less concerned with these things than the philosophers they condemn. Indeed such announcements have repeatedly turned out, over more than two hundred years, to be nothing but striking ways of presenting new (or sometimes, more recently, not so new) theories about the familiar subject-matter. It is perhaps time to take them much less seriously.

Here an objector might admit that, at some level of generality, Aristotle, Locke and Wittgenstein can be said all to have been concerned at least some of the time with the same subject-matter, but claim that the kinds of things said about it, the motives for saying them and the ways of supporting them have changed so radically that the classification of both Aristotle and Wittgenstein as ‘philosophers’ is about as revealing as the classification of both astrology and astronomy as sciences of the stars. Yet this analogy misses a crucial difference. We know, for example, that Aristotle’s account of the names of natural kinds, implicated as it was with a now untenable conception of the relationship between species and individual, can nevertheless influence a present-day philosopher’s thought about the same topic; whereas the astrological reasonings of John Dee have no significance for Roger Penrose. Not that anything written by Aristotle or, for that matter, by Locke could have been written today. Yet suspended in what is now alien or even absurd may be something that can still count as insight and truth. Even Locke’s moral theology, a line of thought which is surely as dead for us as any in philosophy, embodies (or so I will argue) a certain insight into the essential character of morality. Probably the most difficult part of the historiography of philosophy is the task of precipitating such insight without clouding the meaning of the text. It is crucially important that the text be kept at its historical distance even as we explore and try to articulate its contribution to philosophical knowledge in terms acceptable to ourselves. We need to retain a clear understanding of how and why those terms might not (almost inevitably, would not) have been acceptable or even intelligible to the author of the text. It is not a simple matter of separating wheat from chaff. It is perhaps more like identifying the truth in a largely misleading caricature, unfair at every point: a truth, as it is important to stress, of which we might not previously have been aware.

There are more reasons than one for insisting that historical distance be kept, that we remain explicitly and in detail aware of those differences in motive and presupposition and theoretical context which colour everything written by a philosopher like Locke, and which make it virtually impossible that anything he wrote can simply be picked up and taken for truth. One reason is that, in so far as we fail to keep our distance and so conflate explaining what Locke meant with assessing its truth, we inevitably lay ourselves open to the charge that we find insight in the text only because we interpret it anachronistically. The sceptic about the possibility of philosophy can make use of the history of philosophy for his purposes by laying emphasis on whatever is most obviously alien and seems most worthless to us, belonging too clearly to conditions other than our own. The optimist’s reply must not be to deny this otherness, or to rest optimism on the claim that there are doctrines to which it does not extend, and which need only be filleted out of a text like the *Essay* for us to derive from it what is still worth having. We need to show that, *despite* their comprehensive differences from what any philosopher would be prepared to say today, despite, that is to say, what is unacceptable in them, Locke’s arguments have something to reveal to us.

Yet such a demonstration of the continuing value of the past is less difficult than it may seem, since after all the alien is not necessarily worthless. A second and, for philosophers, perhaps more compelling reason for insisting on the otherness of the past lies precisely in the value of exploring a viewpoint different from our own. The interpretation and assessment of philosophical argument are particularly worth keeping apart just *because* philosophy is not like natural science. Both philosophy and science are explanatory and both are theoretical, but the explanations of science, at any rate of developed science, involve commitment to determinate conditional predictions. Hence there are experimental checks which, corrigible or (to use a word borrowed by philosophy from jurisprudence) 'defeasible' as they themselves may be, serve to keep theory and the replacement of theory heading in the broad direction of truth. To set such checks aside is one way of being 'unscientific'. In philosophy there are no such checks, for philosophical explanations are *a priori*. A similar disadvantage appears if we compare philosophy with mathematics. Formal proof is of limited value in philosophy, if only because premises and principles are likely to be as much at issue as conclusions. Formal 'analysis', as a method, at best gives to philosophical argument the nature of an argument as to whether an attempted formalization of a bit of natural language is successful and, if not, why not. The argument lies outside what has been formalized, whereas mathematical reasoning is wholly formal. In epistemology and metaphysics we have in the end only our judgement and 'intuitions' to guide us in distinguishing good explanation from bad. It is therefore important to order our intuitions systematically: formalization may sometimes help, but cannot do so comprehensively.

All this has two consequences. On the one hand, philosophy is an intensely personal activity. Philosophical truth has to be worked for: it is not simply to be accepted as such from acknowledged experts (much as the experts might wish that it were). Expertise cannot be demonstrated by results: in this, philosophical reasoning is like moral reasoning. Philosophers can help one another, but philosophical inquiry cannot successfully be conducted as the kind of corporate enterprise that natural science has become today. On the other hand, philosophical beliefs are exceptionally liable to be determined by such non-rational motives as fashion, loyalty, wishful thinking, intellectual bullying, inflexibility, charisma and institutional politics as well as by insight or, for that matter, honest and intelligent misjudgement. Nor is it surprising if the less fundamental should sometimes determine the more fundamental, so that there can occur the absurdity of a particular moral, political or religious orientation's motivating a theory of knowledge or being.

It is not here denied that effectively irreversible progress is possible in philosophy. That, indeed, occurred in the seventeenth century with respect to philosophical conceptions of the natural world and natural science (although even here ancient materials were to hand.) Yet it is only too easy for progress in one respect to be matched by regression in another, and for progress to be involved with new error, as in the case of seventeenth-century mechanism itself.

Or insight can be pushed too far, so that fashion swings from one paradox to the opposite paradox. Despite the possibility of progress, it seems that philosophy cannot easily cast off the ancient scandal of endless, unresolved dispute.

That is why the history of philosophy is important to philosophers. On the one hand the past is an invaluable repository of unfashionable truth for those with the sympathy, judgement and patience to discern and make use of it. On the other hand its study, with all the objectivity we can muster, is perhaps the only discipline or method by which we can expect to improve our capacity to stand above the non-rational determinants of our own beliefs and concerns, or to call our own deepest presuppositions to account. For the history of philosophy is the story of how philosophy got where it is. Given the nature of philosophy and of the intellectual life, it is incredible that it should have got there either by pure and clear-eyed reason alone, or without some retrievable losses by the way.

How far what follows is a convincing illustration of these optimistic principles I must leave to the reader to judge. I have tried, for reasons which should by now be clear, to separate exposition from philosophical comment, as far as has seemed practicable by placing them in separate sections. In order to bring out what can be gained philosophically from the *Essay*, I have also felt it necessary to set out, in broad terms, some present-day theory about thought, knowledge, meaning, identity and the like. I have done so not only in order to assess the real distance between Locke and twentieth-century philosophy, but also to consider whether the advantage may not sometimes lie with Locke, even (or especially) where the distance is most striking. My treatment of recent theory is necessarily summary, with less of an attempt to be scholarly in its documentation or, perhaps, to be wholly fair than in my treatment of earlier writers. That will, I hope, be acceptable in what is primarily a discussion of a seventeenth-century philosopher. The modern views I criticize are, for the most part, not short of enthusiastic exponents.

The primary, interlocking elements of Locke's system are a general theory of ideas and knowledge, the main topic of the present volume, and a theory of 'substances', 'modes' and relations as constituting the subject-matter of the various sciences, a theory which is examined in the second volume. In the consideration of this system I have been unable to avoid all repetition, for more reasons than one. First, systematic philosophy cannot be neatly parcelled into distinct topics, so that the same move in argument has often needed to be discussed in a variety of contexts. Second, in the exposition of such a wide range of material it has seemed best to include in the earlier discussions a general sketch of Locke's epistemology and its relation to his ontology, and so of most of the themes subsequently filled out in greater detail.

There are, then, several layers of exposition which I have tried to bring into productive and explanatory relation: the exposition of certain antecedents of the *Essay*; the exposition of the main doctrines of the *Essay* (and sometimes of early reactions to those doctrines); the exposition of current theories; and the exposition of my own conclusions. A number of topics have been treated rather

less thoroughly than that, whether out of ignorance or for lack of philosophical ideas. I am acutely aware, for example, that the history of theology and of Natural Law theory lies largely beyond the boundaries of my knowledge, and I have had little to say on my own behalf about the nature of space and time. Nevertheless my procedure has made for a long work. I hope that it is a work which will persuade at least some of its readers that philosophy done without regard to the history of philosophy is two-dimensional, and others, that a discipline which flourished in the seventeenth century survives, if less prestigiously, in our own.

Part I

Ideas

Introduction to Part I

An Essay concerning Human Understanding was a late product of that amazing period in which the Aristotelian view of the natural world, under attack throughout the Renaissance, was finally overthrown and replaced by mechanistic corpuscularianism. In many respects, not only for its sheer best-selling success in advocacy, Locke's rambling, chatty, repetitive, rhetorical masterpiece deserves the seat once generally accorded it beside Newton's *Principia Mathematica* as the culmination of the process by which the 'new philosophy' replaced the old. It also has the considerable advantage, for the philosopher looking to learn from history, that it sets out a version of anti-dogmatic realism constructed before the onset of the idealism or conceptualism which has dominated philosophy, in one form or another, for the last two centuries. It offers us a view from a standpoint that philosophy has otherwise lost. At the same time it illustrates, perhaps more clearly than any other text, how argument between different forms of realism supplied the seedbed on which it was possible for idealism to grow.

The theory of ideas and knowledge presented in the *Essay* apparently sprang from disparate concerns linked by a common attitude: first, a conviction arrived at early in Locke's life, and natural enough in the circumstances, that dogmatic and arbitrary claims to divinely instilled religious and moral knowledge constitute a danger to political stability and order; second, an active interest in medicine and corpuscularian science, pursued under the influence and guidance of such distinguished mentors as Boyle and Sydenham, which evidently confirmed and broadened his hostility towards dogmatism and his respect for experience. His preference for reason over inspiration directed his earliest extended epistemological argument, which dates from the early 1660s and is incorporated in the work now known as *Essays on the Law of Nature*. In this he claimed that knowledge of our duty to God and our fellows lies within the reach of a human intellect employing concepts and premises derived from sense-experience. His anti-dogmatism in natural philosophy, however, came to follow a more sceptical, if similarly 'empiricist' line: the senses give knowledge of no more than the sensible qualities and powers of particular substances, while the intrinsic properties underlying these appearances and powers remain beyond the

reach of our faculties. The ‘corpuscularian’ or atomic theory is no more than the best available hypothesis.

There was no inconsistency between Locke’s epistemological optimism in ethics and political theory and his pessimism as to the possibility of a proper ‘science’ of nature based on the essences of things. Together with his insistence that careful observation and experiment (natural ‘history’) can yield probabilities sufficiently dependable for the direction of action, they constituted his pious thesis that the ‘candle, that is set up in us’, for all its limitations, ‘shines bright enough for all our purposes’. Yet Locke’s first conception of moral knowledge based on empirical premises seems rather different from his later theory of morality as abstract and demonstrable. On what seems to have been his model in his earlier argument, the senses give to reason, first, evidence of design in the world from which we can infer the existence of a creator whose will or law it is our duty to obey, and, second, knowledge of various characteristics of human nature which reveal God’s particular purposes in creating us, and so the content of this Law of Nature.¹ According to the later model, however, we can in the strict sense demonstrate the existence of an eternal, most powerful and wise creator from the premise known intuitively that I myself exist as a thinking thing, while a determinate law follows *a priori* not just for human beings, but for all rational creatures, at least if they are capable of pleasure and pain.² The senses provide the *ideas* or concepts employed in demonstrative morals, including our inadequate idea of God, but not *premises*.

Locke supported both models in turn by an alleged analogy with mathematics. If his conception of moral knowledge underwent change, his conception of mathematical knowledge did so too. On his earlier account, not only are such mathematical notions as those of a line, a plane and a solid drawn from experience, but ‘other common principles and axioms too’ are given to reason by the senses. In 1671, in *Draft A* of the *Essay*, a similar proposal ascribing our knowledge of axioms to ‘constant observation of our senses especially our eyes’ was later overlaid or replaced by a conception of universal knowledge as going no further than our ideas, as hypothetical and, in effect, as *a priori*.³ Locke must be supposed to have taken a clear decision that the empiricism he wanted to advance was essentially concept-empiricism, according to which all our ideas ultimately derive from experience, rather than the stronger, or at least different, view, knowledge-empiricism, according to which all prepositional knowledge is empirical, ultimately based on sensory knowledge.

If Locke’s epistemology underwent such early development, then he can be said to have moved from the sort of empiricism attributable to Epicurus, and expounded previously in the seventeenth century by Pierre Gassendi, to a rather different theory which combined concept-empiricism with some of the features of theories opposed to empiricism. Epicurean epistemology, as it has come down to us and as Gassendi presented and developed it, seems not to have distinguished sharply between the acquisition through the senses of concepts which give meaning to our words, and the acquisition of prepositional knowledge

constituting definitions and axioms capable of serving as the starting-point of reasoning and science. Both seem to spring together from the repeated experience of particulars.⁴ Gassendi accordingly could claim that all our knowledge rests on our sensory knowledge, and that ‘all the evidence and certainty which is derived from a general proposition depends upon that which has been gathered from an induction of particular instances (*singularia*)’.⁵ Locke, on the other hand, by coming in his later thinking to lay more weight on the difference between the acquisition of ideas and the acquisition of knowledge or belief, and between induction and the formation of universal or abstract ideas, was able to account for *a priori* universal knowledge as something distinct from, and independent of, sensory knowledge. He could thus agree with philosophers such as Descartes and the friend of both Descartes and Gassendi, Marin Mersenne, in taking the intuitive understanding of necessary truth as the paradigm of certain knowledge. Like Gassendi, of course, he continued to assign a much more significant role to the senses than the Cartesians did, not only as the source of all ideas, the ‘materials’ and ‘immediate objects’ of knowledge, but also as the source, independent of any reasoning, of ‘sensitive knowledge’ of the existence of particular things. Yet, unlike Gassendi and like Mersenne, he accorded to ‘sensitive knowledge’ the lowest, rather than the highest degree of ‘evidence’.⁶

Moves such as these, apparently made as he picked a careful way through existing doctrine, ensured that Locke’s epistemology did not replicate any earlier position. That is not to say that his sustained attempt to capture what is due to sense and what to reason never rose above judicious eclecticism. Yet in order to understand his thinking we need to identify the various theoretical forces at work in it, forces in the end, no doubt, too various for their product to enjoy full coherence. The term ‘idea’ itself, which has acquired a kind of mythic notoriety in philosophy, supplies a striking illustration of this need. It is chiefly notorious as a gift to the sceptic: since ideas are postulated as the immediate objects of experience and knowledge, it has seemed to follow that they divide us from the reality which they are supposed to represent. But the term is also notorious, especially in Locke’s case, as a knot of unexamined confusions and damaging ambiguities.⁷ This popular reputation is not based on nothing, but it has not in general been based on a clear understanding of the nexus of theoretical motives which determined the meaning of the term ‘idea’ for the philosophers who used it. In [Part I](#) an attempt will be made to achieve such understanding.

First, however, it may be worth saying a little about the traditional charge of ambiguity which does not depend on such contextual analysis and interpretation. Words can, of course, be ambiguous without being perniciously ambiguous. It is hardly a matter for criticism or moralizing⁸ that, like many ordinary words, Locke’s term enjoys ‘type-token’ ambiguity, i.e. the kind of systematic ambiguity which is instantiated in the ambiguity of the question whether Smith and Jones are reading the same book. Just as there is a sense of ‘book’ in which the same book (the ‘type’) can be in many libraries at once, and a sense in which a book (a ‘token’ of the type) cannot be in more than one place at a time, so there

is one sense of 'idea' in which two people can have the same idea, or one person can have the same idea again, and a different but intimately related sense in which shared or recurrent ideas are impossible in principle.

Another harmless systematic ambiguity affecting expressions like 'have an idea' as Locke used them is related to the distinction between dispositions or capacities on the one hand, and occurrences, or the actual exercise of capacities, on the other. With respect to Locke's inquiry 'whence the Understanding may get all the *Ideas* it has',⁹ it is natural and right to interpret the having of an idea dispositionally: to have ideas 'with Names commonly annexed to them'¹⁰ is to have concepts, and to have the concept of X is to have the capacity actually to conceive or think of X. But it seems too that Locke firmly held this actual conscious conceiving to be what is primarily and even most legitimately referred to by talk of ideas' being in the mind or understanding:

But our *Ideas* being nothing, but actual Perceptions in the Mind, which cease to be anything, when there is no perception of them, this *laying up* of our *Ideas* in the Repository of the Memory, signifies no more but this, that the Mind has a Power, in many cases, to revive Perceptions, which it has once had.¹¹

This ambiguity makes the expression 'have an idea of' rather like the ordinary word 'understand', which is used in a 'dispositional' sense when we talk of someone's understanding English, but in an 'occurrent' sense when we say that someone understood a particular utterance.

Many twentieth-century philosophers, however, have supposed that to possess a concept is to have a capacity which is primarily exercised, not in 'perceptions' or conceivings in the stream of conscious thought, but in overt behaviour: characteristically in linguistic behaviour but also, perhaps, in mute sorting behaviour and the like. To such philosophers it has often seemed that Locke's term 'idea' must suffer from pernicious ambiguity just because it is used both for capacities to act overtly and for covert items in consciousness, even for sensations. Yet, since Locke himself would obviously not have accepted the modern behaviouristic account of what it is to have a concept, the charge of ambiguity is hardly appropriate. He believed (and, as we shall see, he believed it for reasons of theory) that conceiving or thinking of X involves having a sensation or sensory image of X. To accuse someone of using a term which is perniciously ambiguous in some respect is to accuse them of confusion at the level of their language, of committing some fallacy of equivocation or of saying nothing clear. Yet, however ambiguous the term 'idea' may be in other respects, in this respect at least Locke employed it deliberately and unequivocally to express a view which is intelligible enough, even if it is now widely held to be mistaken.

Before exploring the topic of Locke's 'imagism', however, we need to achieve a broader basis for our understanding of the notion of an idea by considering some

of the traditional philosophical psychology and ontology which supplied the background to his argument. Chapter 2 will therefore look at some logical theory, with its associated psychology, while chapters 3 and 4 will turn to theories about the processes of cognition. Locke's most famous principle is that all our ideas come from experience, either from sense-perception or from 'reflection' on our own mental activity. Those ideas which may seem either to be creations of our own or to possess a pure and undervived intellectual content antecedent to experience are in fact achieved, according to Locke, by the operations of 'comparing', 'enlarging', 'compounding' and 'abstracting' what is given in sensation or reflection.¹² These operations are in turn explained with the aid of the distinction between simple and complex ideas. Crudely, we can break down what is given in experience into simple parts, and we can put the parts together in new ways.

Locke enjoyed presenting his doctrine as analogous, in the realm of thought, to Epicurean atomism in physics, in particular as atomism was developed by Robert Boyle.¹³ The analogy can be misleading in several respects. First, as we shall see, Locke's simple ideas are not so much parts as aspects of what is presented in experience. Second, Boyle, broadly following Bacon, accepted a firm distinction between the level of descriptive generalization and the level of explanatory theory, between 'natural history' and 'natural philosophy'. An example of the former would be 'Boyle's Law' itself, a generalization about the behaviour of gases which was stimulated and supported by observation and experiment. Theory, on the other hand, consisted for Boyle in the justification, and speculative application to specific phenomena, of the overarching explanatory hypothesis of corpuscles mechanically interacting in the void. It is clear that Locke's psychological 'atomism' was not offered as an explanatory hypothesis, but as a descriptive account of what is present in consciousness and open to view. He explicitly claimed to be applying the 'Historical, plain Method' to our thought, and disclaimed any intention to 'meddle with the Physical Consideration of the Mind'.¹⁴ The natural basis of consciousness he regarded as unknown, and the dispute between materialists and immaterialists as irresolvable.¹⁵ It follows that the analogy he drew between the composition of ideas and the composition of physical particles was in effect a sort of conceit, doing no philosophical work.

The third way in which the analogy may mislead is by drawing our attention away from what gives Locke's compositionism its real philosophical significance, namely its associations with the long history of compositionism, from the time of Plato and Aristotle, in logic and epistemology. In considering the background to Locke's notion of an idea we shall necessarily also be concerned with the background to his distinction between simple and complex ideas, and to his conception of the 'putting together'¹⁶ of ideas of both kinds to form propositional thoughts. It is this conception which makes it appropriate to begin with the 'dry discourse', as Hobbes put it, of logical theory.

Ideas and compositionality in traditional logic

Although the notion of an idea is with good reason chiefly associated with the theory of knowledge, it may be helpful as a preliminary to discuss its place in relation to traditional logic. It is chiefly through logic, at any rate, that Locke's first readers would have been made familiar with the doctrine that the mind forms complexes out of simples. Logic manuals of the kind used by every European undergraduate were typically divided into three parts, corresponding to the increasing complexity of their subject-matter, as they dealt with terms, with propositions formed by bringing terms together, and with syllogisms composed of ordered propositions. A fourth section on method was sometimes added.

Logic seemed to deal at once with language, with thought and with reality. Aristotle himself, in his main discussions of predication, had moved freely between the linguistic and the ontological modes, writing sometimes of the combination of words or expressions in sentences but at other times of the association of things or entities in an act of predication. We are told, for example, that expressions must be combined, and that the combination must at least implicitly include a verb (paradigmatically the copula) before something capable of truth or falsity comes into existence.¹⁷ On the other hand, one thing (an attribute) may be said or predicated or affirmed of another thing, which is the subject (and paradigmatically a substance—'substances' are 'things' in a narrower or stronger sense, comprising men, horses, axes and the like, as well as such stuffs as gold and water.) The Greek word '*kategoria*' can be taken to mean 'predicate', but the famous list of ten categories or 'predicaments' (i.e. in the language of the early English logics: substance, quantity, quality, relation, place, time, posture, habit, action and affection) was regarded as a classification of things or beings or objects of thought, as much as of expressions. The equally famous characterization of 'accidents' as those things which exist 'in' a subject also concerns things as they are designated by words: as we shall see, it is an intuitive ontological criterion with which other, seemingly more linguistic, criteria can be correlated.¹⁸ But there is also for Aristotle a further level to which logic relates, the level of thought, or that which corresponds in the mind to both words and things, and which does not differ from person to person as language may differ between people of different nations.¹⁹

In late Scholastic text-books logic was commonly introduced as being concerned equally with the structure of thought and the structure of speech. The act of mind which corresponds to a term was said to be the 'simple apprehension' of the meaning of the term, or simple 'concept'. 'Judgement' or mental 'affirmation' corresponds to the ordering or combination of terms in the sentence or proposition. Syllogisms, which include 'probable syllogisms', are the expressions in words of 'ratiocination' or mental discourse, consisting in a chain of judgements. In general it was assumed that thought supplies the link between language and reality: terms and propositions relate to reality because they express concepts and judgements. The relation of terms to things, it was commonly stated, is that of being put in their place, or standing for them, in discourse. In these different ways terms are 'signs' both of concepts and things.²⁰

In so far as they discussed it, the logics seem generally to have assumed that truth is a matter of the correspondence between the way in which terms or simple apprehensions are related in propositions or judgements, and the way in which things are related in the world. The affirmation that the wall is white is true if, and only if, the 'accident', whiteness, inheres in the substance, the wall. Another traditional account, which was employed with emphasis by Thomas Hobbes, treated truth as a relational attribute of sentences: a sentence is true provided only that what the predicate names or stands for is the same thing as what the subject names or stands for.²¹ Both accounts are, of course, consonant with the principle that truth and falsity can exist only where there is combination. Some scholastics, however, allowed that 'simple apprehensions' are themselves truth-bearers in that they involve a conformity between concept and thing. At the same time it was argued that a simple apprehension of X cannot be false: if it were false, then it would not conform to X, and so would not be the apprehension of X.²² Echoes of this rather surprising doctrine occurred, as we shall see, in both Descartes' and Locke's approaches to the topic of true and false ideas.

As Aristotelian philosophy came under attack from the advocates of various forms of mechanism, traditional logic caught some of the abuse. Yet it was too impressive a system simply to be rejected. In general it was seen as something needing reform rather than abolition, something to be incorporated into the new view of things. Descartes himself, it is true, had little time for it. The rules for judging well promised by the logicians would be of value, he evidently believed, only if they were useful to the discovery of truth, and he took it that his own method of analysis and synthesis fulfilled this role better than logical rules for ordering terms and propositions. Syllogism he regarded as chiefly useful for sharpening the wits of undergraduates,²³ while the distinction between terms and propositions is irrelevant to his own distinction between simple and complex, and even to his notion of an idea. If in practice he tended to reserve the word 'idea' for non-propositional conceptions of things or their attributes, he nevertheless drew no explicit theoretical distinction between 'ideas' and 'notions' or 'perceptions', which are as often as not, in his usage, propositional. The general explanation of what an idea is, given in the Third Meditation, makes no

reference at all to the term-proposition dichotomy. We are told, on the contrary, that an idea is the representative element which must be involved in any kind of thought or state of consciousness, whether volition, emotion or judgement, and which determines the content of that thought. Its relation to judgement is simple: an idea is the presentation of a content to the mind, and judgement is the act of giving assent to that content. (The same content might have been willed or desired.) Even an idea expressible non-propositionally can present an object which does not exist, and so, as much as a propositional content, provides 'material' for judgement and belief. It can be 'materially' false even if not 'formally' so.²⁴

Descartes' agreement with the traditional view that truth and falsity are properly or primarily attributes of judgements was no more than verbal, since 'judgement', as it should be clear, had a different meaning in his system from the common one. Judgement was not for him a mental act of proposition-making, but an act of will constituting mental assent to what is 'perceived' by the understanding. He saw a strong analogy between reasonable or unreasonable assent and reasonable or unreasonable voluntary action.²⁵ Few have agreed that the analogy is sufficient for his purpose, but the implausible doctrine that belief is always voluntary slots in well enough with his conception of method and his injunction to give assent only where we have clear and distinct perceptions or ideas. It is we ourselves who are to blame for our mistakes, not our faculties or the God who gives them to us. But the present point is that, just because these were the issues which, rather than logic, interested Descartes, his more or less equivalent terms 'notion', 'perception' and 'idea' relate to the content, whether propositional or not, of any of our thoughts. He took that content in abstraction from the question whether we assent to it or not, or indeed from the question whether we assent to it or desire it. Accordingly his conception of simplicity has nothing to do with the distinction between terms and propositions. An idea or notion is simple when it can be grasped or understood by a simple intuition or act of thought. It is therefore unsurprising that his exemplary list of simple notions includes such 'eternal truths' as the 'common notion' that when two things are both identical with a third thing, they are identical with each other.²⁶

Descartes' indifference towards the structures of traditional logic was not, however, typical of his contemporaries. Hobbes and Gassendi both wrote logics incorporating much traditional doctrine, although neither was less opposed than Descartes to Aristotelian metaphysics and science. The authors of the Port Royal *Logic* wrote as convinced Cartesians, yet they admitted that their chief debt lay elsewhere, with Aristotle.²⁷ One of them, Antoine Arnauld, had shortly before collaborated on a *Grammar*. In both these works the logician's concern with grammatical and logical structure was found a place within Cartesian philosophy by something like brute force, in that the Cartesian distinction between perception and judgement was identified with the traditional distinction between the simple apprehension of the meaning of terms (i.e. the having of ideas) and mental affirmation. Whether assent was for Arnauld an act of will is not clear, but

it was evidently his view that (as was usual in traditional logics) a prepositional thought or content comes into existence only with a prepositional attitude. He made no attempt to explain, or even to explain away, the possibility of what was sometimes called the ‘simple apprehension’ of a prepositional content prior to judgement.²⁸ One argument for distinguishing prepositional content from mental affirmation does, however, seem to have been recognized, since the authors of the *Grammar* apparently envisaged dealing with the prepositional content of questions, commands and wishes by the postulation of further species of combinatory act, side by side with judgement, each with its characteristic expression.²⁹ What is certain is that, in the work of Arnauld, Hobbes, Gassendi and most others who used it, the word ‘idea’ was firmly linked to the meaning or import of terms or ‘names’ as opposed to propositions.

Locke’s employment of ‘idea’ was entirely in accordance with this standard usage. In the manuscript generally regarded as the first draft of the *Essay*, he stated that the mind joins and separates ideas ‘by way of affirmation or negation, which when it comes to be expressed in words is called proposition and in this lies all truth and falsehood’.³⁰ Ideas are introduced in the argument of the *Essay* itself as ‘the parts out of which...propositions are made’.³¹ This does not mean that every element of a prepositional thought is an idea, or that every word in a sentence stands for an idea. It is a common mistake to hold Locke to his loose assertion that the ‘use...of Words, is to be Sensible marks of Ideas’,³² when it is clear enough that the words he here had in mind are ‘names’, words (roughly speaking) which can stand as subjects or predicates. He dealt with other words in a chapter on connective ‘particles’, adopting the orthodox view that they express mental actions or ‘intimations’ linking either ideas or whole affirmations. Here he agreed with contemporary logics that ‘*Is* and *Is not* are the general marks of the Mind, affirming or denying’,³³ while he explained the word ‘but’ (moving from the second to the third part of logic) as one of those by which a speaker expresses ‘the dependence of his Thoughts and Reasonings, one upon another’. He was simply echoing many writers again when he asserted that some words, ‘which are not truly, by themselves, the names of any Ideas, are of constant and indispensable use in Language’.³⁴ If his treatment of such words seems perfunctory, that is because the theme of Book III is the confusion due to the misuse of words. He had little more to say about ‘particles’ than that, while their careful use can prevent confusion, the standard explanations of the meanings of some of them mask multiple ambiguities. Yet the chapter uncovers an important part of the traditional framework of his thinking about logic and language.

That framework chiefly reveals itself in the structure of the *Essay* in a very straightforward way. Book II is about what terms or names express, which is that ‘which the Mind can be employ’d about whilst thinking’³⁵ (i.e. roughly, while affirming and reasoning.) Book IV is about affirmations or prepositional thoughts on the one hand, and reasoning and method on the other, although only roughly in that order. Not surprisingly, Locke’s friend Molyneux suggested to him that he should recast his philosophy in the form of a logic, so that it might

gain acceptance in the universities.³⁶ That does not, however, mean that Locke was significantly more traditional in his thinking than other philosophical proponents of the new science. On the contrary, some of the strongest Aristotelian resonances in the *Essay* (as in the writings of many of Locke's contemporaries) can be ascribed to his desire to confront the adversary on common ground, through the radical rereading of some familiar doctrine or piece of technical terminology. The chapter on 'Reason' in Book IV is a vigorous attack on syllogism reminiscent of Descartes, while Locke's own model for *a priori* reasoning involved an anti-formalist reinterpretation of the notion of a 'middle term' as that of an 'intermediate idea'.

The classification of ideas in Book II as 'simple ideas', 'ideas of simple modes', 'ideas of mixed modes', 'ideas of relations' and 'ideas of substances' should be understood as a deliberate attempt to set up a rival doctrine of categories. Significantly, it is drawn up at the level of thought, for Locke meant to emphasize that these are distinctions for us rather than distinctions picked off reality. There is also a significant novelty in the ordering of the new categories, in that 'substance' has been displaced from its primacy by 'simple ideas'. Broadly speaking (although this characterization could mislead) Locke chose to follow the order of knowledge rather than the order of being. It is perhaps a better, if an obvious thing to say, that the new order, placing quality before substance, reflects the relative simplicity of the concepts ordered. Whereas the really fundamental distinction for the traditional doctrine of categories was (for reasons to be considered in Volume II of the present work) between substance and the other categories, for Locke, at least for certain purposes, it was the distinction between simple ideas and the rest which was most important ('simple modes' being simple only by courtesy and in a limited sense). Traditional logics universally gave substantives like 'man' or 'horse' as their paradigms of 'simple terms': the paradigm of complexity arises when substantive is combined with adjective, substance with accident. Locke, on the other hand, had his own reasons, lying in his epistemology, for insisting that our ideas of substances are complex.

It may seem that, since all Lockean ideas, whether simple or complex, are equivalent to the logicians' 'simple apprehensions', the notion of the simple parts of a complex affirmation can have little connection with the notion of the composition of complex ideas out of simple ones. Nevertheless, especially in his early thoughts on the matter, Locke himself felt that there was such a connection. He wrote in 1671 that those simple ideas 'are properly [called] simple apprehensions to which we apply the names that others doe',³⁷ thus both assimilating and distinguishing the notions of a 'simple idea' and of a 'simple apprehension'. (The ground for his treating the latter as a restricted class of the former presumably lay in the traditional explanation of a simple apprehension as the apprehension of the meaning of a term.) More significantly, he saw complexity as prepositional. The first active composition of simple ideas is to form the ideas of specific substances: or, as he himself put it, 'the first

affirmation or negation of our minds are about those material objects in the framing of our Ideas of them'.³⁸ Locke thought that, in forming the complex idea of a species of substance out of simple ideas of sensible qualities, we 'in effect'³⁹ affirm that the qualities do in general exist together in the same substance: 'though the whole compounded idea being knowne under one name and taken altogether considered as one thing as man horse water lead etc. they may be treated of as simple apprehensions' (i.e. as single concepts corresponding to terms).⁴⁰ Perhaps we should here bear in mind the Epicurean tendency, noted above in [chapter 1](#), to run together the acquisition of concepts with the acquisition of definitional truths. Yet Locke was also in effect attacking the Aristotelian view of the names of specific substances as paradigmatic 'simple terms' by means of an Aristotelian conception of complexity. To unpack what is understood by the name is to unpack an essentially prepositional compound.

This suggestion was modified in the *Essay*, as perhaps it had to be as soon as other sorts of complex ideas were admitted into Locke's scheme. On his later view, it is improper to form the idea of a substance unless we have had experience of something possessing all the qualities which we include in the idea. The idea of a centaur is 'fantastical' and, in a certain sense, false. The idea of a substance thus carries a certain informal implication as to the coexistence of the qualities, but does not involve the affirmation of that coexistence. Nevertheless Locke does not seem to have seen his withdrawal from the position that the complexity of a complex idea is the complexity of affirmation as weakening the opposition between his own view that ideas of substances are complex and the Aristotelian characterization of them as simple. For he argued in the *Essay* itself that language misleads the Aristotelian into treating a substance-idea, such as *horse* or *gold*, 'as one simple *Idea*, which indeed is a complication of many *Ideas* together'.⁴¹ Ideas of substances, 'though they are commonly called simple Apprehensions, and the Names of them simple Terms; yet in effect, are complex and compounded'.⁴² Such criticism of the Aristotelians, if considered only in its immediate context, might seem to be no more than a debating point, conflating two different kinds of simplicity. Yet Locke may not wholly have given up his view that ideas of substances are quasi-propositional. He also had in mind other Aristotelian notions of simplicity and complexity only loosely related to the distinction between terms and propositions. These issues will be discussed in what follows, and more comprehensively in Volume II.

Ideas and epistemology before Locke

The Aristotelian soul is the 'form' of the living thing, the principle of life which activates its 'matter'. Although specific forms were supposed to be ontologically simple, they were nevertheless assigned a certain complexity in so far as they contain the forms of higher genera. Consequently *man* was often spoken of as possessing three forms, souls or lives. Barely as a living thing a man shares the lowest kind of life with plants, the life of nutrition, growth and reproduction. As an animal it enjoys sensation, imagination and the power of self-movement. But the distinctive life of the human being is the life of reason and intellect.

Aristotelian epistemology consisted of an account of the co-operation between the distinctively human and the animal forms or souls, combined with an account of a natural correspondence which was postulated between the faculties or mechanisms of sensation and thought, and the objects of sensation and thought. In the perception of something white the sensible form of white exists in the white thing, in the medium between thing and eye, and in the eye itself; but it exists in a different way in each of these subjects. The mode of its existence in the medium is peculiar, in that the air between object and eye, in bearing whiteness, is not itself white. Nevertheless it seemed certain that the 'sensible species' is somehow transmitted *via* the medium, and so exists there *en route* for the organ of sense. It exists in the eye in another peculiar, but more interesting way, namely as an object of vision. It exists there, as the scholastics put it, 'objectively' or 'intentionally'. The sensible forms are transmitted from the sense organs to the 'common sense' and the seat of the imagination or 'phantasia' (fancy), where they can be revived in memory. This account of perception and imagination, originally set out in Aristotle's *De Anima*, later proved capable of combination with fairly sophisticated geometrical and psychological optics by means of the principle that light and colours are transmitted in rays from which the eye selects in the formation of the visual image.

All sensible forms or species, as well as all images or 'phantasms', are particular: what we perceive becomes the object of universal or scientific thought only with the formation by the 'active intellect' of 'intelligible' and universal species or forms in the passive intellect. In this process of abstraction the intellect draws on the objects of imagination. Aristotle himself asserted

roundly that there is no thought without images. As the sensible form or image of whiteness is whiteness itself existing in the organ of sense or imagination, so the intelligible form of anything is that thing existing intentionally in the intellect, i.e. as an object of intellect. It was generally agreed, however, that the active intellect needs no bodily organ, and constitutes that immaterial part of the soul which is immortal.

When the faculties are functioning properly there is ‘conformity’ or congruity between things as they exist in reality and things as they exist in the mind: between what has ‘formal’ or ‘real’ existence and what has ‘objective’ or ‘intentional’ existence. Indeed, on a natural view they are the same things, for when I see the sun, one and the same thing, the sun, exists both in reality and in my mind, both formally and objectively.

Accounts of perception and thought following this general model came under attack as a part of the process by which Aristotelian physics was pressed aside by the new mechanics. On the broad hypothesis that every event in the material world is reducible to motions of matter subject to mechanical laws, it became unacceptable that an ultimate theoretical description of changes in the material organs of sense or imagination should be given in terms of the intentional or objective existence there of what is perceived or imagined (not to speak of ‘all those little images flitting through the air, called “intentional forms”, which so exercise the imagination of philosophers’).⁴³ Descartes did not deny that an image or phantasm of the object exists in the eye or brain, but he offered a reductive explanation of its representative function. Since what occurs in any part of the body can only be the motion of matter, sensible species can be nothing other than mechanical occurrences systematically caused by the perceived objects. Beyond that systematic correspondence, as Descartes himself spelt out, there need be no resemblance or ‘conformity’ between the object and its corporeal ‘image’. The image is simply an event which carries ‘information’ in the technical modern sense, it is not an intentional object or intrinsically intentional state.⁴⁴

A causal correspondence of this type, it should be said, cannot by itself explain representation. If it could, then every effect systematically produced in one thing by another would be ‘of’ its cause in the way in which a sensation or thought is always of something: every cause would be ‘in’ what it affects in the way in which my friend is in my mind when I think of him. But Descartes’ purpose in reducing physical representation in the sense-organ and brain to systematic mechanical effects was neither to explain representation or ‘intentionality’ in general nor to deny that sensations are intrinsically intentional, but to expel intrinsic intentionality from the body and restrict it to the mind or intellect. In sense-perception, on his theory, the intellect somehow so relates to the images in the brain that sensory *ideas*, i.e. conscious sensations of objects, are formed. Ideas have reference to objects essentially: as we have seen, an idea is that element in any act of thought which determines its content. How the intellect is supposed to concern itself with events in the brain or to refer them to objects is left notoriously obscure, but it is in virtue of their role in stimulating ideas of

sense and imagination, and not only in virtue of their origin, that Cartesian corporeal images can be said to represent things.

The restriction of intrinsic intentionality to the modes or states of an immaterial intellect not only kept the material world uncluttered for the purposes of mechanistic physics, but allowed Descartes to pride himself on an account of the mind as something unitary and indivisible. The Aristotelian mind consists, in effect, of a bundle of seemingly separable, if co-operating faculties, some materially based while one, the intellect, is pure form capable of existing alone in separation from the body. For Descartes, on the other hand, the various functions of the mind are not performed by separable faculties, but by one cognitive faculty or power operating in different ways on different objects. Perception, imagination and memory occur when the intellect applies itself to the relevant parts of the brain, while purely intellectual conception occurs when it acts by itself.⁴⁵ Because the subject of all forms of consciousness is immaterial, the mind is indivisible. A notorious corollary of this doctrine is that animals, lacking intellect, are not conscious, despite having corporeal mechanisms corresponding to sense and imagination. These mechanisms, in Descartes' view, give rise to behaviour directly, without the intervention of consciousness.

Not all proponents of the new physics shared Descartes' enthusiasm for the Platonic or Augustinian embodied soul. Hobbes was not alone in holding that the causal chain constituting the process of sense-perception stops in the brain. For Hobbes, since motion can only bring about motion, sensation can only be motion in the part of brain which constitutes the 'organ of sense'. He attempted a heroic mechanistic explanation of the intentionality or directedness of sensation. The motion which is propagated from the object 'to the innermost part of the organ' meets resistance and gives rise to an opposite 'endeavour' or reaction: 'from the reaction, how little so ever the duration of it be, a phantasm or idea hath its being; which, by reason that the endeavour is now outwards, doth always appear as something situate without the organ.'⁴⁶ Hobbes was what may be called an 'imagist': there is, he held, no separate faculty of intellect, and all cognition can be explained in terms of sense and imagination, which he described as 'decaying sense'.⁴⁷ Since there is no idea not derived from sense, our only idea of substance is the idea of matter. The expression 'immaterial substance' is therefore a contradiction in terms.⁴⁸ Hobbes' imagism and his materialism neatly propped each other up.

If the attractions of mechanistic physics brought about changes in philosophical conceptions of the mind of the perceiver, they had an equal effect, of course, on conceptions of the object perceived. In Aristotelian explanations of perception sensible qualities were treated as irreducibly distinct attributes of the object or body. Colour, for example, is a sensible quality proper to one sense which exists in essential relation to sight but which is not reducible to any other attribute. In addition to its qualities, the object possesses an indefinite number of powers to interact with other bodies so as to give rise to observable effects either in itself or in other things. The seventeenth-century opponents of Aristotelianism

saw this multiplicity, for all the efforts of Aristotelian science, as merely unintelligible and without order. In its place they presented the object as something simple in nature, even if mechanically complex. It consists of corpuscles of matter in motion which in turn cause motions in the surrounding matter. Some of these motions, continued in the regular way through any necessary medium and the organs of sense, reach the brain of the perceiver to cause an appropriate idea in the mind. Other effects on the perceiver are less direct, through motions perceived as changes in other objects. The variety of qualities and powers which we attribute to the object really reflects, not a multiplicity in the object, but a multiplicity in the ways in which the object can cause motions in us and so appear to us. The main thought is very clearly expressed by Kenelm Digby, an English contemporary of Descartes supposed by some at the time to be of comparable genius. Not succinct, the passage is nevertheless worth considering for its remarkable connections with Locke's own doctrines, as with other doctrines old and new.

Imagine I have an apple in my hand: the same fruit worketh different effects upon my severall senses: my eye telleth me it is green or red: my nose that it hath a mellow sent: my taste that it is sweet, and my hand that it is cold and weightie. My senses thus affected send messengers to my phantasie with news of the discoveries they have made: and there, all of them make severall and distinct pictures of what entereth by their doors. So that my Reason (which discourseth upon what it findeth in my phantasie) can consider greenesse by it self, or mellownesse, or sweetness, or coldnesse, or any other quality whatsoever, singly and alone by it self, without relation to any other that is painted in me by the same apple: in which, none of these have any distinction at all, but are one and the same substance of the apple, that maketh various and different impressions upon me, according to the various dispositions of my severall senses: ...so that what is but one entire thing in it self, seemeth to be many things in my understanding: whereby...I shall be in danger...to give actual Beings to the quantity, figure, colour, smell, tast and other accidents of the apple, each of them distinct one from another, as also from the substance which they clothe; because I find the notions of them really distinguished (as if they were different Entities) in my mind. And from thence I may inferre, there is no contradiction in nature to have the accidents really severed from one another, and to have them actually subsist without their substance: and such other mistaken subtilties.⁴⁹

It was the Aristotelians (and above all, Francisco Suarez and his followers) who were supposed to have fallen into the trap which Digby describes. His accusation that they believed that 'real accidents' are separable from their substance was perhaps justified by the doctrine of 'sensible species' itself, although the most obviously vulnerable doctrine in this respect was the orthodox interpretation of

the miracle of the Eucharist, endorsed by Suarez but notorious among the New Philosophers, according to which the sensible accidents of the bread remain while the substance is changed.⁵⁰ Yet Digby's fundamental objection was to the subsumption of the ordinary and non-miraculous case under the notion of a substance bearing qualities in some unintelligible way. Such an account seemed quite inadequate as an explanation of the intimate tie between a thing and its attributes. The new philosophy, on the other hand, with its distinction between quantitative mechanical attributes and sense-relative qualities such as colour and smell (Locke's distinction between primary and secondary qualities) offered to solve the mystery of the 'inherence' of accident 'in' substance. If the object is simply matter, i.e. impenetrable extended substance, then the only intrinsic accidents it can sustain are the determinate sizes, shapes and motions of its parts. There is no mystery about the relation between an extended thing and its shape or limits, or between a moving thing and its motion. As Hobbes remarked, this is something that everyone understands. He continued:

And as for the opinion that some may have, that all other accidents are not in their bodies in the same manner that extension, motion, rest, or figure, are in the same; for example, that colour, heat, odour, virtue, vice, and the like, are otherwise in them, and as they say, *inherent*; I desire they would suspend their judgement for the present, and expect a little, till it be found out by ratiocination, whether these very accidents are not also certain motions either of the mind of the perceiver, or of the bodies themselves which are perceived.

The Aristotelian account, Hobbes suggested, treats accidents as though they were parts or ingredients of bodies, and so themselves substantial; whereas an accident ought to be defined as 'the manner by which any body is conceived', or as 'that faculty of any body, by which it works in us a conception of itself'.⁵¹ Robert Boyle made a similar point, distinguishing 'inseparable accidents' of matter ('inseparable, because being extended, and yet finite it is physically impossible, that it should be devoid of some bulk or other, and of some determinate shape or other') from other 'less simple qualities (as colours, tastes and odours) that belong to bodies on their account'. Inseparable accidents he renamed 'moods or primary affections', attacking the scholastic notion that, in addition to these, 'there are in natural bodies store of real qualities, and other real accidents, which are not only no moods of matter, but are real entities distinct from it'.⁵²

In such arguments as these the mechanist hypothesis was justified philosophically in terms of two intimately related points of contrast with Aristotelianism: its intrinsic perspicuity as a piece of ontology, and its power to explain scientifically what is otherwise left unexplained. The latter has received more attention from historians of science, but explanations of the form, 'This stuff puts people to sleep because of its dormitive power', were objects of scorn

not just because they are pseudo-explanations, too easy, undeniable and vacuous to be useful, but also because, if taken as ultimate causal explanations, they entail an ontologically objectionable model of the relation between bodies and their accidents. Accidents are treated like plums stuck in a plum pudding, in effect as real parts rather than as the abstractions they are. Students of seventeenth-century philosophy need to be sharply aware that the question of what makes a perspicuous philosophy of substance and attribute was inextricably entwined with the question of how a world must be in which causality is fully perspicuous. Commentators on Locke (who held, roughly speaking, that we lack an answer to the former because we lack an answer to the latter) have in particular often failed to make that essential connection.

The Aristotelians, it should be said, did offer some account of a quasi-causal, intelligible bond between substance and attribute. They did so in the doctrine of 'predicates' ('species', 'genus', 'difference', 'property' and 'accident')⁵³ which was summarized in every seventeenth-century logic. It was held that the intellect can penetrate to the simple core or essence of the substance, achieving an understanding of a unitary network of attributes necessarily connected to that core and intelligibly flowing from it. First, sustained observation of the kind of thing in question makes it possible to identify those attributes which are possessed by all members of the kind in virtue of their membership to the kind. These are called 'properties', and are distinguished from 'accidents' in a strict sense, namely attributes which are not, or need not be, possessed by every member of the kind, or by any member throughout its existence. The intellect can then on reflection pick out the fundamental property from which the others flow, or by which they can be explained, in order to arrive at a 'real' or 'simple' definition of the essence of the kind.⁵⁴ Such a definition of a 'species' is by 'genus' and 'difference', each genus being similarly definable by reference to the next higher genus until the highest, the category, (i.e. *substance*), is reached. On this model, reflection on the properties of *man* lead to the recognition that its essence is to be a rational animal, while the essence of an animal is to be a living thing with the powers of sensation and self-movement. The 'difference', rationality, explains the other characteristically human structural and dispositional properties. We can grasp that a rational animal is essentially or by nature a language-user, and that hands by nature subserve intelligence, as we can grasp that an axe, given its defining function, must possess a certain structure.

The mechanist critics of this story of course rejected its paradigms of explanation, which are teleological. In what sense does a horse have an identifiable function or end, apart from the use men put it to or God's unknown purpose in creating it? The 'necessity' of a teleological connection seems in any case grossly loose when compared with the tight, quasi-geometrical necessity of mechanics. Moreover, Aristotelian science suffered from an inherent incompleteness, self-confessed in the notion of an 'accident'. The proponents of mechanism nevertheless retained much of the structure and terminology of

traditional doctrine, reinterpreted and redeployed in the service of their new conception of the world.

An issue which divided the mechanists was the question whether the mind can in fact, as well as in principle, penetrate to the simple core or being of things. Descartes was the most influential of the optimists. Like the Aristotelians he assigned a crucial role to the intellect, but he downgraded the role of sense. His purely intellectual ideas do not derive from ideas of sensation, but are innate. He is, of course, famous for having doubted the deliverances of the senses, but it may be helpful to see him less as a philosopher drawn into dogmatic intellectualism by his efforts to escape from the sceptical predicament, than as someone who early in his life decided that he could by pure thought achieve a remarkable piece of knowledge, the principles of a demonstrable mechanics capable of explaining all natural phenomena in the material world. He was then faced with the epistemological problem, not of avoiding scepticism (over which he probably lost little sleep), but of explaining how he had managed to acquire knowledge about the world *a priori*. Unlike Kant's answer to this same general problem, Descartes' was overtly metaphysical: the intellect can spin knowledge from its own entrails because God created it with innate knowledge of essences and of fundamental principles, principles implicitly used in the interpretation of experience. This implicit knowledge can be revealed and made explicit by a proper method of analysis. It can then be employed synthetically in the scientific explanation of the world as we experience it. If, as we proceed, we accept only what is fully intelligible, that of which we have a clear and distinct idea, we shall not fall into error.

A famous argument applied this framework to the object of perception. Descartes pointed out that each and every sensible quality of a piece of wax (including its size, since the wax can appear to expand and contract according to temperature) may be replaced by a different one as circumstances change. His conclusion was that our ordinary idea of the wax as something which survives such alteration is the purely intellectual idea of a thing or substance which is extended and mutable.⁵⁵ Elsewhere he argued that the change proper to an extended substance is mechanical change, i.e. motion and rest subject to intelligible laws. This looks like a very Aristotelian procedure of stripping off accidents to reveal the simple essence, extension, and the properties necessarily connected to the essence; but it was for Descartes the application of a wider method. He was following the principle that (as Spinoza put it in his exposition of Descartes' philosophy) what is complex and not understood should be broken down into what is simple and intelligible and evidently true. As we have seen, what is 'simple' for Descartes is like an axiom, capable of being understood by a simple act of thought. In this it is not unlike the Aristotelian 'simple definition', and there are other echoes of ancient and Scholastic doctrine. For example, combination is a precondition of falsity, although not of truth: 'simple natures are all self-evident and never contain any falsity'. We cannot have a partial or inadequate grasp of a simple nature: otherwise it would not be simple, but a

composite of what we understand and what we do not understand.⁵⁶ The guarantee that such ideas are materially true is precisely their intelligibility, since a God who allowed reason to lead us astray would be guilty of undivine deceit.

If Descartes was the leading epistemological optimist among the New Philosophers, the leading pessimist among his contemporaries was Gassendi, for whom Descartes' method was hardly better than Aristotle's at uncovering the real nature of the thing which lies behind sensible qualities. Gassendi wrote (at any rate, for most of the time) as an imagist with materialist inclinations, but unlike Hobbes he was not a dogmatic metaphysician with pretensions to a demonstrable mechanics. His thought was very overtly influenced by the main post-Aristotelian Hellenistic philosophies, in the first place by Scepticism. Ancient Scepticism had an enormous effect on the shape of early modern philosophy, chiefly through the works of Sextus Empiricus and Cicero's *Academica*, together with the less sympathetic accounts given by Diogenes Laertius and St Augustine. In all these sources, however, scepticism was presented together with a potential antidote by being set against Epicurean and Stoic (or St Augustine's own) counter-arguments. In his earliest work, *Contra Aristotelicos*, Gassendi drew heavily on Sextus himself, but he left it unfinished, perhaps as he came to prefer the more positive, but still significantly undogmatic Epicurean theory. He adopted the main tenets of Epicurus' atomism and hedonism, as well as the sense-based epistemology some elements of which had been further developed by the Stoics in their long-running feud with the Sceptics. This last debate centred on the question whether there is a criterion by which falsehood can be distinguished from truth. The issue raised an echo in Descartes' adoption of the criterion of clear and distinct perception, but the 'clear and evident' impressions on which Epicureans and Stoics founded knowledge were impressions of the senses. Such foundations were taken to reveal less than those postulated by Descartes: as Lucretius remarked, 'The eyes cannot discover the nature of things.'

A concept closely associated with the notion of a criterion of truth was that of a sign, note or mark through which we can have knowledge of things not evident in themselves. A distinction, found also in Aristotle, was drawn between reminiscent signs, which pertain to what is only sometimes not evident, and indicative signs, which pertain to what is by nature never evident. In the case of reminiscent or empirical signs, a connection can be set up in experience between the sign and what it signifies. In the case of indicative signs, we reason to something which is such that, unless it existed, the sign would not exist: e.g. sweating is an indicative sign of invisible pores in the skin, and certain motions of the body indicate the presence of the soul. The Sceptic Sextus accepted the cogency of reminiscent signs as a basis for opinion, but rejected indicative signs as the notion had been advanced by Stoics, Epicureans and other 'dogmatists'. In an extended criticism of Sextus, Gassendi argued that the proof of pores in the skin is acceptable, since it appeals only to principles founded on experience. He made the same claim, for reasons touched on in [chapter 1](#), above, about geometrical

proofs. More relevant to the present topic, however, is his characterization of sensory appearances as signs, and his endorsement of the Epicurean principle that appearances are always true. Even if the same thing presents different appearances in circumstances in which we may suppose it has not itself changed, or appears different to different observers at the same time (as in the standard sceptical example of water appearing hot to one person and cold to another), we can judge that the difference in appearance indicates a difference in the conditions of perception. If we wrongly place the difference in the object, then it is the judgement which is false, and not the appearance. One of Gassendi's Epicurean examples is the changing appearance of an object as it grows more distant. The appearance is not false, he claimed, although a judgement that the object is itself growing smaller or changing shape would be false.⁵⁷

Despite important differences both in content and presentation (some of which have been mentioned in [chapter 1](#) above), Locke's epistemology is extremely like that of Gassendi both in its broad structure and in many details. Some of the dissimilarities themselves appear in terms which suggest a deliberate decision to differ. It is therefore a little surprising that his notebooks and other records do not supply the same evidence of close reading of Gassendi's philosophical works as they do in the case of many other writers.⁵⁸ The internal evidence, however, points to a strong and early influence, direct or indirect, and arguably to an extent that makes it implausible, although it is possible, to attribute the similarities entirely to Locke's having read the same sources as Gassendi. (He did own copies of Lucretius' poem, Cicero's *Academica* and Diogenes Laertius' *De Vitis Philosophorum*, as well as of Gassendi's responses to Descartes' *Meditations*.) Some of the differences and similarities between them will be noted below. The first to be considered, in the section immediately following, is Locke's employment of the ancient thought that sensory ideas are signs or marks and, as such, are always true.

Simple and complex ideas

Locke's conception of the object of experience, and of how far we can know it, is carefully and systematically opposed both to Aristotelian and to Cartesian epistemology. I have considered the latter theories, first with respect to their account of the mental processes of perception and conception, and then with respect to their ontology of the object. In Locke's case I will reverse that order.

Locke's starting-point in his consideration of the object was, unsurprisingly, the same as that of Digby and Descartes. In our experience of a physical object we are faced with a multiplicity of sensible 'accidents', or distinct ways in which the object affects the senses directly or indirectly, each of which, as Digby had put it, can be considered 'singly and alone by itself'. As Locke stated the model (with apparent allusion to Descartes' piece of wax),

Though the Qualities that affect our Senses, are, in the things themselves, so united and blended, that there is no separation, no distance between them; yet 'tis plain, the *Ideas* they produce in the Mind, enter by the Senses simple and unmixed. For though the Sight and Touch often take in from the same Object, at the same time, different Ideas; as a Man sees at once Motion and Colour; the Hand feels Softness and Warmth in the same piece of Wax: Yet the simple *Ideas* thus united in the same Subject, are as perfectly distinct, as those that come in by different Senses.⁵⁹

For Locke, however, we have no way of moving from this multiplicity to a conception of the unitary essence or core of the object as it is in itself. As it was seen in [chapter 2](#) above, he insisted against the Aristotelians that our ideas of substances are all complex. The only idea that we can form of a substance comprises the ideas of the experienced accidents (powers as well as directly sensible qualities) together with the obscure idea of something to which they belong, and which is responsible for their occurrence together. Locke's falling away from his initial view that this complexity is propositional, and for that reason opposed to the simplicity of a 'simple apprehension', was also discussed above. Yet his sense, apparently retained, that he was here operating with something like the Aristotelian notion of complexity had some justification. Even

at its origin that notion had proved capable of breaking free from the context of straightforward predication.

Aristotle himself associated simplicity with the categories: every simple term falls into just one category. Yet that account, as he recognized, leaves open the possibility of a sort of combination of items from different categories to produce, not a proposition, but a new predicate falling into no one category. For reasons to be considered in Volume II, he provided only one kind of example of such a predicate, in which substance is combined with accident as *man* is combined with *musical* to form *musician*. Like any term, ‘musician’ asserts nothing by itself, yet its compound nature shows itself in predication. ‘John is a musician’ means the same as the compound proposition, ‘John is a man and John is musical’. ‘Musician’ is therefore quite different from ‘man’, since (even if we accept the traditional definition) ‘John is a man’ does not fall apart as ‘John is an animal and John is rational’. On the Aristotelian view, that is because ‘Man is a rational animal’ is a ‘real definition’ of the unitary essence or nature of a species. ‘Man’ classifies the individual according to its nature, placing it in a natural kind. Musicians, by contrast, do not as such belong to a natural kind with a genuine essence, but to a class arrived at by arbitrary combination. Correspondingly musicians are not natural individuals *qua* musicians but *qua* human beings.⁶⁰

This doctrine raises issues which are still very important philosophically and which are still not satisfactorily explained. (A full-scale attempt to explain them will be made in Volume II of the present work.) The doctrine was not regularly set out in the standard seventeenth-century logics, and it is not easy to be certain what conscious echoes of it appeared in the writings of opponents of Aristotelianism. The Port Royal *Logic*, however, stated that we have ideas of three sorts of being: things (or substances); modes (or accidents); and things-as-modified. Many of our ideas are of things-as-modified: we commonly know, distinguish and name substances by their accidents rather than by the common natures in which their being really consists.⁶¹ Locke was making a related, if more sceptical point in asserting that all our ideas of substances are complex, and there is some evidence that he was aware of the Aristotelian association. Since we can have no knowledge of the essence or nature of things, we are condemned to ideas of them which are logically composite, combining the idea of an unknown substance with the ideas of its experienced accidents. No definition we can supply for any substance constitutes a ‘real’ definition.

However that may be, the significance Locke placed on the simple-complex distinction seems tailor-made for a confrontation with Descartes. Although he denied the possibility of the intellect’s penetrating to a simple essence of material things, he agreed with Descartes both that the possibility of falsity arises only with complexity, and that simple ideas constitute the dependable link between thought and reality. That is not, of course, because Locke’s paradigmatic simple ideas are innate, but for entirely the opposite reason,

because they are acquired through the senses. Like Descartes, he at one point introduced his doctrine by means of a pretended scepticism:

'Tis evident, the Mind knows not Things immediately, but only by the intervention of the *Ideas* it has of them. *Our Knowledge* therefore is *real*, only so far as there is a conformity between our *Ideas* and the reality of Things. But what shall be here the Criterion? How shall the Mind, when it perceives nothing but its own *Ideas*, know that they agree with Things themselves?⁶²

That the question was not put in any spirit of perplexity is shown by Locke's clear, confident, highly theoretical and anti-Cartesian answer to it. That answer hinges on a neat and very interesting theory of representation.

Lockean simple ideas are, as we have seen, necessarily received through sensation or reflection. To ignore for the present the special case of 'reflection', they are caused in us by external things acting on our sensory mechanisms. For that very reason simple ideas must be taken to correspond to their objects in regular and orderly ways, even if we are ignorant of the nature of those objects and of how they act on us. A simple idea is therefore, as Epicurus and Gassendi had held, a natural sign of its cause. As such it is a 'sign' in another sense too, since it is naturally fitted to represent or 'signify' in thought that feature of real things, whatever it may be, which is in general responsible for our receiving ideas or sensations of that type. So the simple idea of white received in sensation and capable of being recalled in imagination stands in the natural language of thought for whatever in the object underlies or constitutes its general power regularly to cause just that sensation in us, the power Locke called the 'quality' of the object. The epistemological sign is also the semantic sign of this quality.

It follows that, in the terms of Locke's pseudo-sceptical question, simple ideas are all necessarily 'real' and 'conform to' or 'agree with' things themselves. They cannot be 'fantastical':

simple Ideas are not fictions of our Fancies, but the natural and regular productions of Things without us, really operating upon us; and so carry with them all the conformity which is intended.... Thus the *Idea* of Whiteness, or Bitterness, as it is in the Mind, exactly answering that Power which is in any Body to produce it there, has all the real conformity it can, or ought to have, with Things without us.⁶³

Similar arguments go to prove that simple ideas are necessarily adequate and also 'true' in the sense in which ideas can be said to be true if something exists conforming to them. They are 'Marks of Distinction in Things, whereby we may be able to discern one Thing from another.' Locke echoes the ancient characterization of an 'indicative' sign as the sign of something non-evident which is a necessary condition of the sign: they are 'such Appearances, as are

produced in us, and must be suitable to those Powers he has placed in external objects, or else they could not be produced in us'.⁶⁴

It is crucial to the opposition between the two philosophers that the very ideas which for Locke epitomize simplicity, ideas of colours, tastes, heat and cold and the like, were taken by Descartes to exemplify composition. Yet they started from much the same understanding of our cognitive relationship to such attributes. As Descartes put it,

when we say that we perceive colours in objects, this is really just the same as saying that we perceive something in the objects whose nature we do not know, but which produces in us a certain very clear and vivid sensation which we call the sensation of colour.⁶⁵

This combination, in the idea of a thing's colour, of what is clearly known with what is unknown ensured, for Descartes, that the idea is composite and so liable to falsity. In the *Meditations* he linked the possibility of such an idea's being 'materially' false with the possibility that (as he supposed to be the case with the idea of cold) it represents a mere privation, or nothing, as something real and positive. Since all sensory ideas present their objects as positive qualities, there is no telling from their intrinsic character which is true and which is false. Descartes' position (against which Arnauld brought the Scholastic argument mentioned above that simple apprehensions must in the nature of reference and representation be true) seems to have been that, given that cold is a privation, the sensory idea of cold can only be false.⁶⁶ In the *Principles of Philosophy*, however, he set aside the argument from the possibility of privative causes and in effect allowed for two sorts of 'obscure and confused' ideas of a colour: if we cautiously judge that the external cause of our sensation is unknown, then we avoid error; but if (as it is natural and normal for us to do from childhood) we rashly judge that what is in the object is qualitatively like the sensation, then our judgement is erroneous. The implication is that there are two possible sensory ideas of colour-in-the-object, one materially true and the other materially false. Both are complex.⁶⁷

On Locke's account, on the other hand, the sensation or image *is* the idea of the colour, and his criterion for its simplicity is phenomenal, the limits of phenomenal discrimination: each simple idea, being 'in it self uncompounded, contains in it nothing but *one uniform Appearance*'.⁶⁸ With phenomenal simplicity goes indefinability: the idea of a colour cannot be conveyed in words. But that was a familiar point not in contention. It had already been taken by Leibniz, who dismissed it as irrelevant to the Cartesian position:

We cannot explain what red is to the blind; nor can we make any such object clear to others unless by bringing them into the presence of the thing and making them see, smell or taste it.... It is nevertheless certain that the

notions of these qualities are composite, and may be resolved, in as much as they have their various causes (*quippe cum causas suas habeant*).⁶⁹

The real disagreement here between Locke and the Cartesians concerned the role of the causal relation between idea and object. For Locke this relation *constitutes* the basic representative relation: it determines what the idea represents but does not enter into the content of the idea in such a way as to make the idea complex. The simple appearance is taken by the mind as the sign of its unknown cause, but the mind has no choice in the matter since that is what a natural sign signifies. Speculations as to the intrinsic nature of its cause, whether true or false, are irrelevant to the signification of the idea or to its truth. Simple ideas do not, he claimed, ‘become liable to any Imputation of *Falsehood*, if the Mind (as in most Men I believe it does) judges these Ideas to be in the Things themselves’. For they are ‘as real distinguishing characters, whether they be only constant Effects, or else exact Resemblances of something in the things themselves.’⁷⁰ Locke wanted to say that we can distinguish things and have knowledge of their existence through sense-perception, but that no method of analysis of the ideas we form of them will bring us to knowledge of their underlying nature. Analysis of complex ideas into simple ones will serve only to make explicit the various simple representative (i.e. direct causal) relationships which are given as such in sense-experience. Upon these our assumption that our thought concerns the world must stand, but can stand safely. Descartes’ conception of the role of analysis involved a sharp and (he supposed) intrinsically evident line between our clear and distinct intellectual apprehension of extension and its modes and the obscure and confused ideas of merely sensory qualities. If this line is not in fact clear to us, that is because of mental confusion. For Locke, in contrast, ‘sensitive knowledge’ is even-handed as between primary and secondary qualities, and the distinction between them comes later as a speculative hypothesis which for us can be at best (as he thought it is) highly probable. The content of the hypothesis, that some of our simple ideas represent their causes as they are in things themselves, or that there is in some cases a ‘resemblance’ between the idea and the external quality represented, will shortly be examined.

No doubt because of the argument of the *Meditations*, Locke opened his chapter on primary and secondary qualities with the topic of ideas of privations, such as the ideas of cold, darkness, black and rest. He saw an inquiry into ‘the nature of things existing without us’ as simply irrelevant to the status of ‘the *Idea*, as it is in the Understanding’. The possibility that a ‘clear and positive’ idea should represent a ‘privative cause’ is evidently not disturbing, since it is easy to speculate that

all Sensation being produced in us, only by different degrees and modes of Motions in our animal Spirits, variously agitated by external Objects, the

abatement of any former motion, must as necessarily produce a new sensation, as the variation or increase of it; and so introduce a new *Idea*.

In any case, the notion of a privation is suspect, since it is open to dispute 'whether Rest be any more a privation than Motion'.⁷¹ The whole argument is a little rambling, and even less sharply anti-Cartesian, perhaps, than the present summary of it. Its real purpose seems to have been to preface the topic of primary and secondary qualities with a hard distinction between an inquiry into our ideas as they are in the mind and an inquiry into the nature of their causes. The primary and secondary distinction is a part of the second, not of the first. It is a 'little Excursion into Natural Philosophy', physical speculation rather than descriptive psychology. All this constituted, in effect, a rejection of the way Descartes had seen the distinction and our knowledge of it. Yet, as we shall find, it brought its own problems.

It can be said, then, that a fundamental difference of view over the nature of what ties thought to reality, whether innate structural and interpretive principles or reliable experiential building-blocks, found expression through the medium of compositionism, i.e. through the notions of simple and complex and the doctrine that only the complex can be false. If we recognize the epistemological point of that dispute, we can perhaps avoid exaggerating both the limitations which compositionism placed on Locke's thought and apparent inconsistencies in his argument. Take, for example, the problem raised by the question, 'What are the simple ideas of extension?' Here Locke was pulled in two directions. He seems to have seen that the general or determinable concept of extension (which is the Cartesian 'simple notion') has its claims. Yet, put to the question, he unsurprisingly preferred what he called a 'sensible Point', 'the least portion of Space and Extension, whereof we have a clear and distinct *Idea*'. He did not deny that such a point has extension, and so parts, and he in effect conceded that his paradigms of simple and complex do not fit extension very neatly, but he was no more in real retreat than Leibniz had been over the indefinability of colours. As a footnote to the fifth edition of the *Essay* reported him,

If the *Idea* of Extension is so peculiar, that it cannot exactly agree with the Definition that he has given of those *Simple Ideas*, so that it differs in some manner from all others of that kind, he thinks 'tis better to leave it there expos'd to this Difficulty, than to make a new Division in his Favour. 'Tis enough for Mr *Locke* that his Meaning can be understood.

The objection is dismissed as a pedantic nicety, since it does not touch the main point that our power of forming, for example, ideas of geometrical figures falling *outside* our experience, can be explained as a capacity to build out of what has been given *in* experience. Nor does the objection undermine the claim that we know in 'actual sensation' that the ideas of extension which are received correspond systematically to external causes.⁷²

In another passage⁷³ Locke seems to have been ready to allow that, after all, no idea whatsoever enjoys absolute simplicity, since all ideas, 'when attentively considered', include 'some kind of relation' in them. 'And sensible Qualities,' he asked, 'what are they but the powers of different Bodies, in relation to our Perception.... And if considered in the things themselves, do they not depend on the Bulk, Figure, Texture and Motion of the Parts?' Given its universality, the point is obviously not an abject capitulation to the Cartesian view that ideas of colours and the like are complex, nor is it an equally abject flight from compositionism. It is a part of an argument for including ideas of powers in the class of simple ideas as elements of complex ideas, in particular complex ideas of substances. Locke had in mind that in forming our idea of a causal or dispositional attribute, e.g. the power of wax to be melted by moderate heat, we employ the regularly observed effect to identify and stand for whatever in the object causes it. Indeed the same regular effect conjoined with the idea of power may represent something in the 'agent', the fire which has the power to melt the wax, as well as something in the 'patient', the wax which is liable to be melted. He saw this as analogous to the way in which the sensation or idea of white adequately represents whatever in the object causes that sensation in us, and sufficiently analogous to justify extension of the notion of simplicity to ideas of powers. There are in fact difficulties for this move even within the terms of Locke's argument, since he failed to notice the possibility of 'false' and 'fantastical' ideas of powers: e.g. of the imaginary power possessed by the imaginary 'philosophers' stone' of transmuting lead into gold. Yet, although this objection does in a way touch the nub of Locke's interest in simplicity, it does not clearly demonstrate that the analogy between ideas of powers and ideas of sensible qualities is insufficient for his purposes. Both, after all, on his account have reference to aspects of things which are only accessible to us through their effects and which are not, as such, capable of further analysis. Both, moreover, seem to serve as constituents of our complex ideas of substances. It is, as we shall see, in the theory of substance that the fertility of the analogy is revealed.

5

Ideas as images

As it has been suggested above, there was an important difference of opinion, fairly long-standing by the time Locke wrote, as to the nature of the processes involved in the formation of our concepts of the objects of experience. The dispute hinged on the question of the existence of a distinct faculty of intellect. For the Aristotelians, the active intellect was a separate, immaterial, immortal faculty necessary for the formation of universal notions. For Descartes, the pure intellect was simply the unitary immaterial soul as it thinks on its own, without attending to the common sense or corporeal imagination. Hobbes and Gassendi, on the other hand, rejected an intellect in either sense (although Gassendi wavered in his last work when faced with the task of allowing for immortality).⁷⁴ Hobbes was at any rate not alone in holding firmly that for any thought, particular or universal, ‘we need no other faculty than that of our imagination’.⁷⁵ Since the seat of the imagination was widely agreed to be material, the issue was commonly seen as ontological as well as epistemological. Nevertheless there could be a purely epistemological motive for imagism, and some imagists, such as Joseph Glanvill, the author of *The Vanity of Dogmatizing*, were firm believers in an immaterial soul.

The question whether Locke’s ‘ideas’ are all sensory images (when not actual sensations) is simply the historical question of where he stood on these issues. It is a question which has yet to be settled by modern commentators, few of whom even approach it from this direction. Despite the relative unpopularity of an affirmative answer, the grounds for holding him an imagist are conclusive.

First, it should be said that whether Locke is taken to have rejected or to have accepted purely intellectual, non-sensory concepts, his motive for doing the one or the other must be supposed to have been epistemological. The ontological or ‘physical’ question is not only set aside throughout the *Essay* as irrelevant to its purpose, but is officially pronounced incapable of resolution.⁷⁶ If we have a certain natural inclination towards dualism, that, he suggested, is only because we do not ‘apprehend’ how in general ‘the Operations of Mind, viz. Thinking, Reasoning, Fearing, etc.... can belong to Body, or be produced by it’.⁷⁷ Elsewhere he emphasized how poor a reason our own incomprehension is for any conclusion at all, since many things undoubtedly occur which we do not

understand.⁷⁸ Yet even when most favourably disposed towards dualism, he made no appeal to a distinction between imagination and intellect, or to the existence of non-sensory ideas.

Locke explained the word 'idea' in a single, much-quoted sentence:

It being that Term which, I think, serves best to stand for whatsoever is the Object of the Understanding when a Man thinks, I have used it to express whatever is meant by *Phantasm*, *Notion*, *Species*, or whatever it is, which the Mind can be employ'd about in thinking.⁷⁹

This account has often been taken as the expression of an intention to employ the term generically, for intellectual concepts (i.e. 'notions') as well as for images and sense-impressions. Such a reading is improbable, for Locke's remark seems to have been modelled on a number of similar remarks by Gassendi. Here is one of them:

An image...is what revolves before the mind and is as it were presented to it when we are thinking of anything. It gets given a number of other names. It is also called an 'idea' and a 'species',... a 'notion', 'forenotion' or 'anticipation',...and again a 'concept'. Another name is 'phantasm', in that it is located in the fancy or imaginative faculty. I will more frequently call it an 'idea', because that is now a familiar and well-worn term, and suffers less from ambiguity than the others.⁸⁰

For an imagist, the only thing 'which the Mind can be employ'd about in thinking' is a sensation or image, and so that is all there is for a 'notion' to be. Gassendi's explanation illustrates how the term 'idea' had come to have a sort of theory-neutral use (rather than a broad or generic one) despite the fact that its chief proponent, Descartes, had chosen it, as he stated, because its Scholastic employment for God's concepts of his creatures associated it clearly with intellect rather than with sense. It may be that Locke was simply following the fashion in preferring the word, but he may also have felt that 'image' and 'phantasm' were too closely tied to the standard account of the imagination as a material organ. Moreover 'image', as Gassendi suggested and as we shall see, was certainly ambiguous.

'Image', 'imagination', 'fancy' and 'intellect' are words Locke used relatively seldom in the *Essay*, except when 'imagine' and 'fancy' are used as untheoretical synonyms for 'suppose' as in catch-phrases such as 'I imagine' or 'I fancy'. Two contexts in which 'image' occurs, however, deserve to be noted. In one, it is equivalent to 'resemblance', or to 'representation' in a special sense in which it connotes 'resemblance' between idea and object. Bearing secondary qualities in mind, Locke pointed out that we should not assume that the ideas or perceptions caused by bodies are 'exactly the Images and Resemblances of something existing without us'⁸¹ since, as he later remarked, they are not 'all of

them the Images or Representations of what does exist'.⁸² This metaphorical usage need not now concern us, but the other context is more directly relevant. In the chapter on the 'retention' of ideas, Locke speculated that the constitution of our bodies (in the fourth edition adding, 'and the make of our animal spirits') influences memory, 'since we oftentimes find a Disease quite strip the Mind of all its Ideas, and the flames of a Fever...calcine all those Images to dust and confusion, which seemed to be as lasting, as if in Marble'.⁸³ It would be very difficult to read this chapter without taking recalled sensory ideas to be at least mental images, if not corporeal ones. Take, for example, the final appeal to birds' powers of mimicry on behalf of the view that animals too have the faculty of 'laying up their ideas' (not to speak of having them): 'Birds can approach their Notes nearer and nearer by degrees to a Tune play'd yesterday'.⁸⁴ The proposal is that yesterday's tune is present for the purpose of comparison. There is no suggestion in this whole account of retention, moreover, that it is not entirely comprehensive. There is no hint of the Cartesian doctrine of a purely intellectual memory as well as one that involves the imagination. Yet the power of retention is an important part of what makes thought possible for Locke, in the sense in which 'thought' is opposed to sensation. We are told that the child 'thinks more, the more it has matter to think on' as it 'comes, by degrees...to retain and distinguish the Ideas the Senses convey to it'.⁸⁵ Even more roundly, 'if anyone will say, there are *Ideas* in the mind, that are not in the memory; I desire him to explain himself, and make what he says intelligible'.⁸⁶ Given these connections, an imagist account of memory entails an imagist account of thought in general.

One might even be tempted to conclude from the chapter on retention that Locke held a materialist theory of memory and thought, for the account is very reminiscent of Hobbes' view of memory and imagination as 'decaying sense'. Such an interpretation would, of course, be in conflict with Locke's official agnosticism on the issue of materialism, and elsewhere he explicitly denies that we know the role of 'fleeting Animal Spirits' or corporeal organs in memory.⁸⁷ Nevertheless he seems at any rate not to have been averse to the physiological overtones of the word 'image', and the same goes for 'fancy' and 'imagination'. A relevant passage is the discussion of the association of ideas in the fourth edition. Like others, he was inclined to attribute association to the 'Trains of Motion in the Animal Spirits'.⁸⁸ In another late addition he attributed religious enthusiasm to 'the ungrounded Fancies of a Man's own Brain'.⁸⁹ Elsewhere he asked, 'Is there any thing so extravagant, as the Imaginations of Mens Brains?'.⁹⁰ All these passages are consonant with Cartesian physiological explanations of error (especially prominent in Malebranche's *Search after Truth*) which attribute irrationality and delusion to the influence of the corporeal imagination. Where Locke differed radically from the Cartesians, however, was in making no move towards ascribing rationality and right-thinking to some higher, non-material faculty of pure intellect operating with pure, non-sensory ideas. So far from doing that, he was prepared to describe even irrational associations of ideas as

‘intellectual habits’.⁹¹ Similarly, he asserted that ‘the first Capacity of Humane Intellect, is, That the mind is fitted to receive the Impressions made on it’ through sense and reflection.⁹² ‘Intellect’, like ‘understanding’, had for him a use no more specific than ‘mind’: *Draft A* was written ‘*de Intellects Humano*’, and his working title remained *De Intellectu*.

It may be a little surprising that a seventeenth-century philosopher who was an imagist should have refrained from the explicit assertion that thought can be explained adequately in terms of the imagination without recourse to a faculty of pure intellect. Yet it would be far more surprising if a philosopher on the other side, who believed in intellectual notions as well as sensory images, should have refrained from discussing or even mentioning the difference and relation between the faculties of intellect and imagination. Locke did offer explanations of various mental faculties and disabilities, wit, judgement, stupidity, invention, sagacity, madness and so forth, but all were explained in terms of what their possessor can or cannot do with ideas, or of the order in which ideas arise in his mind, never in terms of different sorts of idea. Moreover it is easy to find passages in which Locke was at least implicitly engaging in the dispute as to the need to postulate pure intellect over and above sense and imagination. One, for example, is evidently a response to Descartes’ discussion of just that issue in the Sixth Meditation. Descartes had tried to prove the difference between conceiving and imagining by arguing that a chiliagon may be clearly conceived, but cannot be imagined. In imagination ‘I may construct in my mind a confused representation of some figure; but it is clear that this is not a chiliagon. For it differs in no way from the representation I should form if I were thinking of...any figure with very many sides.’ Locke’s response was to deny that we can form a clear and distinct idea of the *shape* of (in his case) a chiliagon. We can reason about it accurately because we have a precise idea of the *number* of its sides.⁹³ Such an idea is possible, as Locke had gone to some pains to show in the chapter on number, because of the technique of counting, which breaks the apprehension of large numbers into easy and recorded steps.⁹⁴ This explanation is very like Hobbes’,⁹⁵ and the whole argument is evidently a rebuttal of the Cartesian distinction between imagination and intellect from an imagist point of view. Much the same seems true of his rejection of a positive or complete idea of infinity.⁹⁶ His motive is not obscure: he wished to deny that in intellect we have a faculty by which we can penetrate to a reality beyond appearances, or to a conception of reality different in kind from our experience of it. Hobbes’ imagism was part and parcel of his materialism, Gassendi’s tied in both with his materialist inclinations and his scepticism, but Locke’s was motivated above all by his scepticism or agnosticism about the essences of things.

Hostility to Descartes’ conception of intellect pervades the *Essay*. The rejection of innate ideas is one part of it, even if Descartes was not the only target here, but there is also Locke’s careful reinterpretation of the Cartesian notions of clarity and distinctness, the truth and falsity of ideas, and so forth. For

Descartes a clear and distinct idea or perception is one which involves intellectual understanding. For Locke, by contrast, ideas are clear

when they are such as the Objects themselves, from whence they were taken, did or might, in a well-ordered Sensation or Perception, present them. Whilst the Memory retains them thus, and can produce them to the Mind, when-ever it has occasion to consider them, they are *clear Ideas*.⁹⁷

In this definition Locke refused to advance a step beyond sense and memory or imagination. The definition of distinctness is more complex, but it involves the relation between ideas and words (themselves sensible sounds and marks), not a separate faculty of intellect.

Another example is supplied by the difference between their conceptions of simplicity, already discussed in [chapter 4](#), above. For Descartes simplicity was logical or conceptual. Locke in effect offered to explain conceptual simplicity as phenomenal simplicity, at the same time as he ran conceptions together with sensory appearances: a simple idea ‘contains in it nothing but *one uniform Appearance*, or Conception in the mind, and is not distinguishable into different *Ideas*’.⁹⁸

A popular line of thought runs like this: Locke believed that we have ideas of such things as jealousy and a lie, not to speak of the ideas of unity, being, cause, power and substance. It is not possible to have *images* of these things. Therefore at least some of Locke’s ideas were intended not to be images but concepts of some more refined type. This argument begs the question, and at best, if its premises proved anything, they would prove that imagists who accept (as all did) that we have any of these concepts allow an impossibility. If that is what Locke was guilty of, he was in the distinguished company of Hobbes and Gassendi, not to speak of their ancient antecedents. Yet in any case the argument ignores Locke’s actual treatment of the ideas in question. Those mentioned form a miscellaneous group which he approached, whether successfully or not, in a variety of ways, but each explanation is palpably consonant with their being ideas of imagination. Briefly, jealousy and a lie involve ideas of reflection, which Locke carefully explained as ‘very like’ sense: it ‘might properly be call’d internal Sense.’⁹⁹ For the purposes of the general theory, except in relatively minor ways, reflection is thereafter assumed to be as much a sense as sight, and ideas of reflection to be in effect a class of retained and revivable impressions which for the most part need no special treatment. The idea of substance, however, is explained in quite another way. In so far as it is not the image of something, that is not because it is a special non-sensory idea, but because it is without determinate content, an ‘obscure’ and ‘relative’ place-marker for the unknown. It is simply the idea of ‘something’ somehow responsible for the union of attributes perceived together. It is at the limits of abstraction, involving, as he said to Stillingfleet, ‘the general idea of something, or being’, the only idea we have by which we can think of whatever is unknown.¹⁰⁰

What then is it, on Locke's view, to have the general idea of being before our mind? The answer to that must depend on our understanding of his doctrine of abstract ideas. As it will be argued below, for Locke an abstract idea is a particular perception or image 'partially considered',¹⁰¹ as he put it, and given a certain function in thought. Thus we have the abstract idea of existence before us when we consider either the idea in our mind 'as being actually being there', or external things 'as being actually without us; which is that they exist, or have *Existence*'.¹⁰² The abstract idea of existence, in other words, is the idea of anything considered barely as existing. As for unity, 'whatever we can consider as one thing, whether a real Being, or *Idea*, suggests to the Understanding, the *Idea* of unity'. Units are not mysterious abstract entities, but anything whatsoever, 'Men, Angels, Actions, Thoughts',¹⁰³ considered *as* units. Thus it is not possible to have the abstract idea of *two* in mind without having in mind the idea of some dual in particular, but considered barely *as* a dual. (Ideas of large numbers, as we have seen, are explained in terms of the technique of counting.)

Perhaps the most fundamental and conclusive argument for taking ideas as images, however, concerns the development of Locke's theory of universal knowledge. To pursue it very far now would unduly anticipate later discussions, but the gist can be given. The doctrine of abstraction in the *Essay* is above all an imagist theory of the *a priori* perception or intuition of universal truth, a theory very like those of Hobbes and Berkeley. As for most seventeenth-century philosophers, geometry supplied Locke with his chief paradigms of *a priori* knowledge and understanding. In the geometrical context the issue between imagism and intellectualism took the form of a dispute about the role of diagrams. On the Cartesian view, diagrams, actual or imagined, at best prompt the intellect to form or make explicit its purely mental conceptions. For imagists like Locke, however, actual or imagined diagrams are themselves the essential objects of geometrical reasoning. Actual diagrams have certain advantages over imagined ones, as Locke explained:

Diagrams drawn on Paper are Copies of the *Ideas* in the Mind, and not liable to the Uncertainty that Words carry in their Signification. An Angle, Circle, or Square, drawn in Lines, lies open to the view, and cannot be mistaken: It remains unchangeable, and may at leisure be considered, and examined, and the Demonstration be revised, and all the parts of it may be gone over more than once, without any danger of the least change in the *Ideas*.¹⁰⁴

What Descartes called 'the botched constructions of the imagination'¹⁰⁵ are here contrasted, not with the absolute permanence of the intelligible, but with the relative permanence of the sensible.

This point helps to explain an otherwise mysterious argument against scepticism of the senses:

And though mathematical demonstrations depend not upon sense, yet the examining them by Diagrams, gives great credit to the Evidence of our Sight, and seems to give it a Certainty approaching to that of the Demonstration it self. For it would be very strange, that a Man should allow it for an undeniable Truth, that two Angles of a Figure, which he measures by Lines and Angles of a Diagram, should be bigger one than the other; and yet doubt of the Existence of those Lines and Angles, which by looking on, he makes use of to measure that by.¹⁰⁶

For Locke, we can perceive geometrical relations between particular ideas (e.g. the equality of certain angles) literally with our eyes, and this knowledge becomes universal in so far as we employ those ideas as representative members of classes defined by some precise, abstractly considered point of resemblance, e.g. the figure's having three sides. As he wrote,

the immediate Object of all our Reasoning and Knowledge, is nothing but Particulars. Every Man's Reasoning and Knowledge, is only about the *Ideas* existing in his own Mind, which are truly, every one of them, particular Existences: and our Knowledge and Reasoning about other Things, is only as they correspond with those our particular *Ideas*. So that the Perception of the Agreement, or Disagreement of our particular *Ideas*, is the whole and utmost of all our Knowledge. Universality is but accidental to it, and consists only in this, That the particular *Ideas*, about which it is, are such, as more than one particular Thing can correspond with, and be represented by.¹⁰⁷

How *good* a theory this is may be left for later consideration. The present concern is with *what* it is.

To sum up: Locke's 'ideas', when not occurring in actual sensation (or 'reflection'), are sensory images or quasi-sensations. They are *mental* images in that to have an (occurrent) idea is evidently for Locke to be in a state of consciousness: an idea must be 'taken notice of' in order to exist. On the question of the relation between ideas (whether sensations or mental images) and events in the brain (or what Descartes called corporeal images) Locke was officially agnostic. He did not rule out the possibility that they are identical, and he seems to have believed that physical mechanisms are comprehensively involved in memory and imagination, as in sense-perception. The motive for his imagism was, however, epistemological rather than ontological or materialist. We conceive of things as we experience them, and our capacity to speculate about what lies beyond our experience is limited to the materials supplied by experience. This limitation does not prevent us from achieving *a priori* universal knowledge, as in geometry, despite the traditional principle that the senses and the imagination deal only in particulars. For the particulars in which they deal may be employed as universal representatives or signs.

Ideas as intentional acts and ideas as intentional objects

There is evidence that Locke took an interest in a quite different dispute about the nature of ideas, between Malebranche and Arnauld, which erupted during the period over which he was working on the *Essay*. He explicitly criticized Malebranche's side of the argument, but it is a problem of interpretation how far he was (or saw himself to be) in agreement with Arnauld, and with the Cartesian doctrine to which Arnauld appeals. It is worth examining this doctrine with some care, since it raises some profound questions about intentionality, or the relation of thought to its objects.

Descartes' fundamental conception of an 'idea', as was said in [chapter 2](#) above, was of an intrinsically representational mode of thought. In virtue of their representing function all ideas are like 'images of things', although Descartes did not, of course, mean by this metaphor that all ideas are sensory images involving the corporeal imagination. In the argument for God's existence which follows this explanation he drew the Scholastic distinction between the 'formal' reality or being of an idea and its 'objective' reality. The argument hinges on the traditional notion of a scale of reality or being, with God at the top followed in order by finite spirits, corporeal substances and, finally, modifications of these substances. Another premise is the principle that everything on this scale is the effect of something at least no lower on the scale than itself. The crucial question is, where on this scale does the (allegedly) innate idea of a perfect being fall? Judged by its 'formal reality' it has the lowly status of a non-substance, a mode of thought; but it also possesses 'objective reality', i.e. the level of being of its object, which gives it the highest status of all. It must therefore have been caused by a perfect being.¹⁰⁸

In defence of this argument Descartes glossed the claim that the idea possesses or contains two sorts of reality by saying that the word 'idea' is ambiguous: it can be taken either 'for an operation of the intellect... or objectively, for the thing represented by that operation'.¹⁰⁹ The suggestion that there is a sense in which the 'idea' is 'the thing which is thought of in so far as it has objective being in the intellect' was pursued in another defence of the argument: the idea of the sun was said to be the sun itself existing in the mind. Existing 'objectively' was contrasted with existing 'formally' or (as Descartes put it)

existing 'as the sun does in the heavens'. Descartes' critic Caterus complained that 'objective' existence is no sort of existence, since what is 'in my mind' in this sense need not exist at all, and so requires no cause. If an *existing* object comes to be 'in my mind', the only intrinsic difference is in my thought, not in the object. Descartes, however, argued that we must allow to the sun in so far as it is thought of a special mode of being or existence which, although inferior to formal existence, is 'not merely nothing'. We therefore need to suppose an adequate cause, not only for every object which exists in reality, but for every object which we can conceive of.¹¹⁰

Caterus was surely right. The claim that 'intentional objects' (as they are now most commonly called) have a special sort of being is ontologically intolerable. Yet it has a certain charm from a logical point of view. If it is possible to talk about such an object as the dagger which appeared to Macbeth, it may be supposed that the dagger must be there to be talked about, even if it lacks full-blooded existence. So, by drawing a picture of a horse, even of no horse in particular, it may seem that we have thought the horse into a thin kind of being, existence in a picture. Russell objected that this will commit us to granting existence to objects whose description involves self-contradiction, since logically impossible objects can be represented. But there are other problems. Suppose that it is said that a weak sense of 'exists' operates within the intentional context of the picture (or any other representation), while a strong sense relates the intentional object of the picture to the world. It should immediately be clear that both 'senses', contrary to the hypothesis, can be employed within the intentional context: an artist may depict a real horse looking at a painted horse, and Macbeth's dagger is an imaginary object among the really existent objects in Shakespeare's play. So if a horse in a picture can be said to have a yellow mane just because that is how it is depicted, can't we also say that it really exists, since that (let us suppose) is how it is depicted too? Descartes characterized the *real* existence of the sun as existence 'in the sky', but of course in my mind too the sun exists in the sky. There seems, then, to be some truth in Hume's view that to imagine a horse is to imagine a horse existing. In other words, the phenomenon which prompted Descartes to multiply modes of existence (and prompts others even now to multiply senses of 'exist') is better managed by recognizing an ambiguity of scope: 'The horse in the picture exists' does not mean the same as 'The horse exists in the picture' (or 'In the picture, the horse exists'), but that is not because 'exists' is in any way ambiguous. None of this, it is true, supplies an answer to a general problem about intentionality: how is it possible to suppose, feign, hallucinate or discuss what does not exist? How can any representation have non-existent intentional objects? The present point is only that the way to attack this problem is not by a proliferation of modes of existence or meanings of 'exist'.

Even if we grant Descartes his notion of 'objective' or intentional existence, or perhaps just because of that notion, it is not easy to understand his view of the relationship between the intentional object and the real object, the sun existing in my mind and the sun existing in the sky. He agreed with Catenas that when I

come to think of the sun, all that happens to the real sun is that it acquires an 'extrinsic denomination', or purely relational attribute. That admission seems to require that the sun on which he bestows 'objective being' cannot be the real sun, since he denies that objective being can be reduced to an extrinsic relation. Yet could Descartes really have held that the intentional object of our thought of God is not the real God? It seems obviously preferable to hold that the distinction between them is not a real distinction, but only a distinction of thought. When we distinguish the man in the picture from the sitter and assign them different attributes (one has a straight nose, the other a crooked one), the true objects of comparison are not different individuals, but how an individual is represented and how he really is. Unless, despite their differences, the sun as it exists in my mind were identical with the sun as it exists in the heavens, I could never think of the real sun, even to misconceive of it, and, in general, thoughts and representations would be confined to intentional objects, as opposed to real objects. The question arises, no doubt, under what conditions intentional objects are identical with real ones. That question is the complement of the question asked at the end of the previous paragraph, for it is the question of how any representation can have *existent* intentional objects. The two questions can and should be answered together, by an examination of the conditions under which representations succeed in having, or fail to have, reference to things.

The dispute between Malebranche and Arnauld seems at least in part to have flowed from the uncertainties in Descartes' position. Malebranche's doctrine that we perceive all things in God was one of the more exotic seventeenth-century epistemologies. At its heart is a neoplatonic theory of universals. The extension which is the object of geometrical reasoning and knowledge is, he held, neither real extension (or body) nor a mode of thought, but an entity in its own right, 'intelligible extension', with which the mind is united in intellectual perception. It has the status assigned to universals by much medieval philosophy and, in effect, by Descartes: it is a perfection of God, the archetype in God's mind of created extension. The same model is applied to sense-perception. Perceived *ideas* are sharply distinguished from *sensations* and feelings, which are simply modifications of the mind, and it is made equally clear that the idea of the sun and the real sun are entirely distinct beings: the sun we immediately perceive is intimately united to the soul, whereas the real sun is millions of miles away. The former could exist even if the latter did not. The former changes size when the latter does not. The former is full of light, whereas there is no sensible light outside the mind, only motion. Nevertheless it is through our union with the former, which (stripped of its merely sensory elements) is an idea in God, that we can have knowledge of the latter, which God has created according to his idea. In effect Malebranche drove a wedge between the intentional object and the real object (as Descartes himself was inclined to do), giving the intentional object an ontological status independent of finite minds.¹¹¹

Arnauld's reaction was to deny the existence of any such 'representative beings', i.e. independent intentional objects which represent real objects (and

which he compared to the ‘intentional species’ of the Aristotelians). He rejected the principle that the object of thought must be present to the mind in any but the tautological sense in which, simply by being thought of, an absent friend is ‘present to’ or ‘in’ our minds. Objective presence does not require local presence, and it is mere confusion to suppose that distance prevents the sun from being objectively in the mind.

On Arnauld’s view an idea or perception is a thought of something, a modification of the mind or soul. It can, however, be regarded in either of two relations: to the soul which is modified, or to the thing perceived as it is objectively in the soul. Corresponding to Descartes’ two senses of ‘idea’, Arnauld proposed different meanings for ‘idea’ and ‘perception’: ‘perception’ should be used more specifically for the soul’s perceivings considered as acts of the mind, while ‘idea’ should be used when what is chiefly in question is the object of perception as it is in the soul.¹¹² Here the argument becomes confused, since he held both that the term ‘idea’ denotes the object as it is in the soul (the intentional object) and that idea and perception are not two different entities, but the same thing regarded in two ways. Thus a similar problem arises as for Descartes with respect to the relationship between the intentional object and the real object. If the idea or intentional object is not distinct from the perception, it is difficult to see how it could ever be identical with a real object. Yet if what is ‘in the mind’ or ‘present to the mind’ is the idea and not the real object, then the notion of being present to the mind can hardly be illustrated, as Arnauld illustrated it, by the ordinary thought that a friend can be continually before our mind. Indeed Arnauld drew a distinction between the *immediate* object of perception, ‘the objective reality of the sun, which is present to my mind’, and the *mediate* object, ‘the possible or actual sun which is outside my mind’. Where he differed from Malebranche was in holding that the former is nothing but our own intentional state considered with respect to its content. He even offered this as an explanation of the reflexivity of consciousness: our awareness of objects is necessarily a form of reflexive awareness, a cognitive state which, in this special way, has itself as its object because it has a thing-in-the-mind as its object.¹¹³ It is not surprising that Malebranche found Arnauld’s argument inherently sceptical, or that Reid later found it inconsistent.

The question of the relationship between Locke’s conception of an idea and the theories of Arnauld and Malebranche has recently been mooted by several commentators on Locke, in particular by John Yolton.¹¹⁴ Yolton’s purpose is to defend Lockean ideas against the traditional criticism that, as the immediate objects of perception and thought, they stand as a barrier between the mind and external things. On the interpretation Yolton rejects, since we do not perceive external objects, our knowledge of them must depend on a doubtful causal inference from the ideas we do perceive. Such an interpretation would be appropriate, he thinks, only if Locke had treated an idea as a thing in its own right, an entity distinct both from the perceiver and from the external object which it represents. Malebranche did regard ideas in this way, but the kind of theory from

which Yolton seems most eager to differentiate Locke's is one which assimilates 'ideas' to the 'sense-data' or blank 'qualia' of twentieth-century empiricism. For Malebranche, 'intelligible extension' and other ideas have a special ontological status as perfections of God, but they are still essentially intentional objects, 'representative beings' existing only in the mind, even if independently of finite minds. The immediately perceived colour-patches and the like of the logical empiricism of fifty years ago, on the other hand, are more like Hume's 'impressions': they exist in no second-class way, they really exist, and indeed they comprise most, if not all of the things we can know to exist.

Yolton plausibly suggests that modern commentators looking back on Locke have too readily assumed that 'ideas' are like 'sense-data'. Even Thomas Reid, in 1785, whose criticism of Locke¹¹⁵ is in many respects a classic statement of the interpretation which Yolton rejects, was looking back through Hume for the origin of the errors which led to Hume's scepticism. In Yolton's view, on the other hand, Locke followed Descartes in using the term 'idea' with conscious, harmless ambiguity (the ambiguity Arnauld avoided by distinguishing 'perception' from 'idea') and with a proper ontological parsimony. He used the term *either* for the thought of an object *or* for the object thought of (as it is thought of), but not for a third thing *between* thought and object. On this interpretation the notorious principle shared by both Arnauld and Locke that (as Arnauld stated it) 'we have no knowledge of what is outside us except through the mediation of the ideas within us'¹¹⁶ would seem conveniently to dissolve into an innocuous tautology: we can only have knowledge of what is represented in our thought, as it is represented. When Locke wrote 'Every man's Reasoning and Knowledge, is only about the Ideas existing in his own Mind',¹¹⁷ he may have meant no more than that we can only think and reason about things as we conceive of them. He may not have meant that we are condemned to think about a peculiar kind of objects of which physical objects are the supposed causes.

It does seem right that 'idea' is employed in the *Essay* with effectively the Cartesian ambiguity. Ideas are often identified with perceptions and less often with sensations, and these, Yolton contends, are modifications of the mind. A difficulty here is that 'perception' and even 'sensation' might raise the same problem of interpretation as 'idea', but occasionally their status is clear. Pleasure and pain, for example, are unequivocally said to be not only simple ideas, but also 'constitutions of the mind'.¹¹⁸ As for the other Cartesian meaning of 'idea', if we consider, for example, Locke's account of knowledge as the perception of a relation between ideas, it is overwhelmingly attractive to take ideas here to be, not modifications or 'constitutions' of the mind, but the intentional objects of such modifications. Geometrical relations are not relations between thoughts, but relations between such things as triangles and squares, the objects of thoughts. What we contemplate in doing geometry is not the act of conceiving a triangle, but the triangle conceived. Locke's account of the situation is considerably complicated by the theory of abstraction summarized in [chapter 5](#), above. For, as we have seen, the triangle immediately conceived is a particular, fully

determinate triangle which is selectively considered in some respect (*ex hypothesi*, its triangularity) and which represents all such other particulars as resemble it in that respect. He used the term ‘abstract idea’ (or ‘general idea’) sometimes for the determinate particular so employed, sometimes for the selected aspect of the particular and at least once, in a notorious but deliberately paradoxical passage, rather as if our having the ‘abstract idea’ involved the actual presentation to the mind of all the particulars represented by it. The general idea of a triangle ‘must be neither Oblique, nor Rectangle, neither Equilateral, Equicrural, nor Scalenon; but all and none of these at once’.¹¹⁹ The paradox seems to be that these various triangles are *all* of them intentional objects of the thought, although *none* of them is so in particular. The ‘idea’, in the sense of a thought or modification of the mind, is *of* all and none of these at once, but Locke heightens the paradox by taking the term in its alternative sense, so that the general idea is whatever is represented, a sort of comprehensive, but at the same time exclusive, intentional object.

Another context in which it can often be helpful to remember that an ‘idea’ may be an intentional object is the topic of primary and secondary qualities. Take this tortuous piece of self-explanation:

Thus a Snow-ball having the power to produce in us the *Ideas* of White, Cold and Round, the Powers to produce those *Ideas* in us, as they [*scilicet* the powers, not the ideas] are in the Snow-ball, I call *Qualities*; and as they [*scil.* the powers, not the qualities] are Sensations, or Perceptions in our Understandings, I call them *Ideas*.¹²⁰

The passage seems to be saying that what are produced in us by the powers or qualities in question are those powers or qualities themselves ‘as they are...in our Understandings’, in which status they are called ‘ideas’: in other words, ideas are the intentional objects produced in the mind by real objects. (That Locke wrote ‘as they are *Sensations*, or *Perceptions* in our Understandings’ might be taken to endorse the suspicion that these terms are ambiguous too, being capable of meaning *intentional objects*.) Note that, contrary to Caterus, it is not always absurd to say that something in reality ‘produces’ an intentional object or phenomenon, since rainbows are produced by light reflected off raindrops. In the relevant sense rainbows exist only ‘in us’ or ‘in the mind’, for all their physical causes. In a world without vision, there would be no rainbows. Locke held a similar view about colour itself: as a phenomenon it ceases to exist when it is not seen. Take away the senses, ‘and all Colours, Tastes, Odors, and Sounds, as they are such particular *Ideas* [i.e. *qua* intentional objects], vanish and cease and are reduced to their Causes’.¹²¹ This line of thought is, at any rate, not unreasonable.

The most obvious benefit conferred by Yolton’s interpretation, with respect to the topic of primary and secondary qualities, relates to the very principle of the distinction between them: i.e. the principle that ‘the *Ideas of primary Qualities* of

Bodies, *are Resemblances* of them,... but the *Ideas, produced* in us by these *Secondary Qualities, have no resemblance* of them at all'.¹²² For example, a 'Circle or Square are the same, whether in *Idea* or Existence; in the Mind or in the *Manna*'.¹²³ If ideas are like 'sense-data', this claim gives rise to the problem (at least) of understanding how what is not in space can have spatial attributes. But if ideas are intentional objects, the comparison lies between the object as represented in sense-experience and the object as it is. To find difficulty with the notion of an idea which is extended or, indeed, enormous or solid would be like jibbing at talk of a cheerfully smiling picture on the ground that a piece of canvas has neither face nor emotions.

There is another broad consideration which *prima facie* supports Yolton's thesis. Although Locke was a sceptical philosopher, he was not sceptical about the existence of things, but about their essence. The barrier to knowledge seems to have stood, for him, not between sense-experience and external objects, but between things as they appear and things as they are. He often seems to have been more prepared than Descartes or Arnauld to grant 'ideas' a kind of external existence, as if they constitute a world, the world as we experience it. The passage on the snowball and its qualities, quoted just above, continues: 'which *Ideas*, if I speak of sometimes, as in the things themselves, I would be understood to mean those Qualities in the Objects which produce them in us'. As he was prepared to see the idea as the quality existing in the mind, so he often refers to the quality as the idea existing in the object. The more it appears that Locke seriously intended to identify intentional object and real object, the more his epistemology may seem to be a form of 'direct realism' eschewing intermediaries between thought and things. Nevertheless there are serious obstacles to that interpretation of it.

Ideas as natural signs

Despite the attractions of an interpretation of Locke's usage of 'idea' in the light of Descartes' and Arnauld's explanation of its ambiguity, the identification of ideas (in one sense of 'idea') with intentional objects supplies us with a difficult and slippery interpretive model. Even Arnauld, as we have seen, found it impossible to maintain what at one point he seems to promise: a view which allows that we immediately perceive external things. His reason seems to have been that, because an account of the intentional object of a perception is an account of the perception, giving what Descartes had called its 'form', it is therefore not an account of the real object. Such an account can be true, after all, even when no real object exists; and if its truth or falsity is to that extent unaffected by what lies outside us, how can it be *concerned* with what lies outside us? It might be said that we can make an actually existing thing the logical subject of the account, in saying, for example, that the (real) sun appears red: but then (so Descartes and, even more clearly, Arnauld insisted) what is predicated of the real thing is a merely relational attribute, whereas what is predicated of the sun in the mind is intrinsic to it. The real sun and the sun in the mind are distinct. So even if Locke used 'idea' consistently to mean the intentional object of thought, he might still have had something like Arnauld's reflexive awareness in mind when he wrote that '*the Mind*, in all its Thoughts and Reasonings, hath no other immediate Object but its own *Ideas*, which alone it does or can contemplate'.¹²⁴ Whether that would be 'direct realism' is doubtful.

However that may be, there are more serious difficulties for Yolton's interpretation. Locke's splendid demolition of Malebranche's claim to have supplied the only intelligible or explanatory account of perception contains no clear signs of a theoretical affinity with Arnauld.¹²⁵ In the *Essay*, on the other hand, we find passages which seem to have been designed expressly in order to *distance* its doctrine from Arnauld's. Arnauld attributed Malebranche's errors to a false principle: 'None of the bodies which our soul knows can be present to it in themselves. They must be present to it by means of images that represent them.'¹²⁶ If Locke read this general rebuke, as we can be sure he did, he remained unrepentant: 'For since the Things, the Mind contemplates, are none of them, besides it self, present to the Understanding, 'tis necessary that something

else, as a Sign or Representation of the thing it considers, should be present to it: And these are *Ideas*.¹²⁷ It is clear that, unless presence to the mind is here a relation something like Arnauld's 'local presence', a relation with a stronger ontological significance than Arnauld's 'objective presence', Locke's argument does not get off the ground. How could anyone hold that what the mind contemplates is not 'objectively present' to it? Moreover his characterization of ideas as signs identifies them as the constituents of mental propositions:¹²⁸ i.e. as the ideas perceived to 'agree' or 'disagree' in prepositional knowledge. That is why, in this same section, Locke said that they are made 'use of for the understanding of Things'. Yet it is just this function of ideas, as was seen in [chapter 6](#) above, which supplies one of the best arguments for Yolton's view that by 'ideas' Locke often meant intentional objects. It is therefore difficult to avoid the conclusion that Locke was saying that immediate cognitive presence requires quasi-local, ontic presence, and that the requirement was of some theoretical importance to him. The immediate objects of my thought must be with me, where I am.

The only other interpretive response to this argument which would seem to have the slightest chance of being right is a view of it as a mere slip or muddle due, as it may be, to the malign influence of traditional epistemology. According to the Aristotelian psychology of knowledge, considered in [chapter 2](#) above, when I see or imagine a colour the sensible form or species exists 'in me' in *both* Arnauld's senses. The intentional existence of the form of red in the organ of sense or imagination is a special way in which red exists *physically* in the organ, a way of existing there which is different from its way of existing in the object, and different again from its way of existing in the medium. Malebranche's theory is evidence that mechanistic physics and the demise of forms did not automatically force apart the two ways of being 'in the mind'. Locke may certainly seem sometimes to have reverted to the discarded model of the transmitted form, as when he speaks of 'these Organs, or the Nerves which are the Conduits, to convey [ideas] from without to their Audience in the Brain, the Mind's Presence-room'.¹²⁹ Such passages are no doubt, as in effect he tells us, consciously figurative and unserious: 'when I say the senses convey [ideas] into the mind, I mean, they from external Objects convey into the mind what produces there those *Perceptions*'.¹³⁰ But was Locke perhaps just occasionally motivated by the feeling that, since forms are not truly transmitted from external thing to brain, then the true object of perception cannot be the external thing or any of its attributes, but must be the end-product of the causal process of perception serving as a mere representative of what initiated it? Can the argument for the necessity of a 'sign' therefore be passed over as no more than such an atavistic and temporary backsliding from the direct realism which (according to Yolton) is Locke's more considered doctrine?

The suggestion is easy to resist. Locke's conception of ideas as present signs of absent significata was not an aberration but part and parcel of his central and very theoretical account, examined in [chapter 4](#) above, of the role of simple ideas

as the natural signs of their regular causes. In virtue of that doctrine we can say that Locke conceived of ideas as the elements in a natural language of thought. The names of a conventional language owe their signification to an arbitrary relation, but ideas are linked to what they signify by a natural, causal relation. That is why the perception of a relation between ideas can constitute ‘real knowledge’ of their significata, whereas the contemplation of mere words gives no instruction as to anything else.

As it has in effect been explained in [chapter 4](#), Locke adopted a purely causal understanding of the ‘conformity’ between ideas and things at least partly in order that ideas of secondary qualities should be beyond reproach as signs. Nevertheless this account of conformity is general and, as we shall see, intimately related to Locke’s explanation of the possibility of ‘sensitive knowledge’. Sensitive knowledge consists in our immediate awareness of the external origin of ideas of sense, and so extends only so far as knowledge of the powers of things.¹³¹ The tenor of this whole approach is to portray ideas as blank effects in us: i.e. physically, ontologically or quasi-locally in us, in us as modifications. Such effects represent external things, at any rate as far as perceptual knowledge is concerned, solely in virtue of their evident external origin. It is just because all our simple ideas are acquired in sense-experience that they point beyond themselves. On this official causal theory, or so it seems, what is given in sensation is the existence and coexistence of powers of things to affect our senses, but nothing else about those things, not even that they are disposed in space. Some effects in us have a spatial character, and we may have a general inclination to believe of all sensory effects that their external causes resemble them, but the distinction between primary and secondary qualities, as Locke presented it, is an overwhelmingly convincing hypothesis posterior to sensitive knowledge rather than something given or known at the moment of sense-perception.

Here a supporter of the Yolton thesis may say that Locke cannot really have meant what he seems to have meant. He must have meant, or must in charity be taken to have meant, that ideas are intrinsically representative, presenting a world to us (or, on the alternative meaning of ‘idea’, *constituting* a presented world) which is ‘in us’ cognitively but not ontologically. What the senses tell us, on this interpretation, is when there are real objects corresponding to the intentional objects or appearances presented: what they do not tell us is whether the real objects resemble the intentional objects.

It would indeed be charitable to understand ‘resemblance’ between psychological effect and material cause as resemblance between the intentional object and the real object, i.e. between the material thing as presented and the same thing as it is in itself. Yet that is not a natural interpretation of all the passages already considered, or of several others in which Locke seems too evidently to have been in the grip of the ‘blank effect’ model. One group of such passages arises in the presentation of the distinction between primary and secondary qualities itself, above all in the analogy between secondary qualities

and pain. Locke, unlike Descartes, hardly ever betrayed an inclination to see the state of pain as intrinsically intentional or representational: i.e. as an unpleasant presentation or appearance of a part of the body. The location of pains, like the distinction between bodily and mental pains, is explained in terms of their causality.¹³² Yet he returned several times to the theme that ideas of secondary qualities are related to their objects in the same way as pains are related to their causes: most tersely, ‘Why is Whiteness and Coldness in Snow, and Pain not, when it produces the one and the other Idea in us...?’¹³³ If sensations are intrinsically intentional that question is easily answered: whiteness and coldness are aspects of snow as it appears to us, whereas pain (in the sense in which pain is an intentional object rather than an intentional state) is an aspect of our body as it appears to us. A pain in my knee (as will be explained in [chapter 23](#), below) is a bit like a secondary quality of my knee, whatever other objects, such as a hammer or knife, may have caused it. But for Locke the rhetorical question seems to be unanswerable just because he is adopting the standpoint of his causal theory of representation.

A related theme has to do with the meaning of the names of sensible qualities. We ‘denominate’ bodies ‘from’ our ideas, which seems to mean that words like ‘blue’ and ‘sweet’ are in the first instance names of ideas, and that they have come to be applied to things by a sort of extension of their primary usage. In the case of secondary qualities the result is ambiguity which, Locke suggested, since it is unrecognized, is improper:

’Tis True, the Things producing in us these simple *Ideas*, are but few of them denominated by us, as if they were only the causes of them; but as if those *Ideas* were real Beings in them. For though Fire be call’d painful to the Touch, whereby is signified the power of producing in us the *Idea* of Pain; yet it is denominated also Light, and Hot; as if Light and Heat, were really something in the Fire, more than a power to excite these *Ideas* in us; and therefore are called *Qualities* in, or of the Fire.¹³⁴

Note that the ambiguity is supposed to extend even to the word ‘quality’. In effect, Locke was attributing the Aristotelians’ ‘real qualities’ to their taking too seriously certain ‘ways of speaking’ which are ‘accommodated to the vulgar Notions’ but which, as he went on to say, despite the misleading ambiguity ‘truly signify’ nothing but powers. It is perhaps worth noting that the argument, which is advanced as a concession within the context of his claim that all our simple ideas are adequate, is less than fully coherent. What he evidently wanted to say is that, while the name of a secondary quality is misleadingly ambiguous, yet, since it stands for a simple idea, when applied to external things it can only signify the natural *significatum* of that idea; i.e. whatever in external things regularly causes that idea. Yet his argument implies that we can employ the names of *primary* qualities for the real attributes of things in precisely the *same* sense as we employ them for ideas, and in doing so speak truly. How, then, can Locke maintain that

the Aristotelians and the vulgar do not manage to say something false, and think something inadequate, whenever they say that grass is green without recognizing the alleged ambiguity of 'green'?

However that may be, the question more relevant to our present concern is whether the allegation of ambiguity can be interpreted in terms of the distinction between intentional objects and real objects. It is not easy to do so. Locke is not, for example, claiming (as some still claim, whether truly or falsely) that words like 'heat', 'light' and 'blue' have an everyday, sense-relative meaning for the phenomenon in question and another, theoretical meaning for the underlying causal basis of the phenomenon (the 'real' as opposed to the 'intentional' object). Both the meanings he identifies are for him entirely ordinary and untheoretical: one is for the idea in us, the other for the external object in virtue of its causing that idea. But what makes it seem certain that Locke is here employing the 'blank effect' model is the role of the analogy with pain. It is an unavoidable conclusion that the 'sensations' or 'simple ideas' produced in us, like the 'sense-data' or 'raw feels' of some twentieth-century empiricism, comprise for Locke what is 'hot' or 'blue' or 'sweet' or the like in the primary sense of these terms. In its secondary sense 'blue' means *having the power to produce an idea which is blue* (in the primary sense of 'blue'). The notion of 'resemblance' between ideas and qualities, and the remark that 'A Circle or Square are the same, whether in *Idea* or Existence', now take on a disturbing ambivalence, being open both to the 'charitable' interpretation and to one which is considerably more problematic.

There is another passage in which the 'blank effect' model seems to have had truly disastrous consequences, and that is the discussion of Molyneux's problem: if a man born blind were made to see, could he distinguish a globe from a cube by sight before he touched them? Locke quoted Molyneux's answer with approval:

Not. For though he has obtain'd the experience of, how a Globe, how a Cube affects his touch; yet he has not yet attained the Experience, that what affects his touch so or so, must affect his sight so or so.¹³⁵

Whatever the reason for Locke's approval, the argument he approves seems to depend on the assumption that all that the visual idea of a shape has in common with a tactual idea of the same shape is its cause: whereas if we suppose that they are intrinsically intentional, what they have in common is their object (i.e. in the other sense of 'idea', they are the *same* idea). Molyneux's argument then, of course, collapses, as it should. On this interpretation Locke's treatment of ideas of two senses here is inconsistent with his assertion elsewhere that spatial ideas are received 'both by seeing and feeling'.¹³⁶ Yet even the 'blank effect' model, operative though it seems to have been, would not have given Locke the assumption he needed. As the idea of a primary quality, the tactual idea of squareness will 'resemble' squareness in the object. Since the same goes for the

visual idea, the two ideas will resemble each other, at any rate with respect to their squareness. So Molyneux's argument fails whether 'ideas' are regarded as intentional objects or as blank effects—unless, of course, its proponent, like Berkeley, is prepared to drop Locke's thesis that ideas of primary qualities are effects which resemble their causes.

It can perhaps help us to understand Locke's approval of Molyneux if we consider the context to which he chose to attach his discussion of the problem when he added it to the second edition.¹³⁷ The section had previously been concerned only with the perception of the depth of three-dimensional objects by sight. Locke agreed with the long-accepted view that in order to form an idea of depth the mind has to improve on what is recorded through the eye. Like Descartes, he postulated a visual datum which corresponds to the immediate retinal stimulus: 'the *Idea* we receive from [a globe] is only a Plain variously colour'd, as is evident in Painting'. The general point, an ancient one in optics, is that the light stimulating the eye cannot in itself distinguish a three-dimensional source from a flat one. Consequently it was supposed that the optical stimulus must be enriched by further information, and that three-dimensional vision (or, more strictly, judgement) occurs as the result of a process of reasoning. One of Descartes' proposals was that the mind takes into account the angles of the eyes focusing on the object and forms a judgement of distance 'by a sort of innate geometrical knowledge shared by all men'.¹³⁸ By this means we can arrive at the sizes and shapes of three-dimensional objects from the sizes and shapes of two-dimensional visual images. We confusedly suppose that the sensation itself is of three-dimensional objects because the geometrical inference has become so customary as to be unnoticed. Locke in effect took over the framework of this explanation, including the odd (if, in the seventeenth century, widely accepted) claim that, strictly speaking, we do not *see* things three-dimensionally. Where he differed from Descartes was over the type of reasoning which is supposed to be employed. Whereas for Descartes the mind reasons at least in part geometrically, drawing on its innate understanding of extension, for Locke its inference is solely based on past experience of 'what alterations are made in the reflections of Light, by the difference of the sensible Figures of Bodies'.¹³⁹ The judgement, 'by an habitual custom, alters the Appearances into their Causes'. As Descartes and earlier optical writers had held, the customary inference is unnoticed, like the immediate inference from speech to its meaning.

It is this account of our customary association between sight and touch which leads naturally enough into the discussion of Molyneux's apparently more radical suggestion as to how far we are 'beholding to experience, improvement and acquired notions'. The whole line of argument was a gift to Berkeley, who proceeded with considerable skill through a quasi-Lockean, if much elaborated treatment of visual depth and Molyneux's problem to a sweeping rejection of the distinction between primary and secondary qualities. But for the present purpose we may notice how well the optical notion of a primitive sensory datum subject to interpretation seems to fit with Locke's causal theory of representation and his

conception of 'sensitive knowledge'. The fit is not perfect, as we have sufficiently seen, but it is at least not surprising that the 'blank effect' model for a simple idea should have seemed a natural one in the optical context.

Ideas: a summary

The argument so far has been an attempt to map at least roughly the theoretical forces which lay behind, and gave point to, Locke's usage of the term 'idea'. Some of the interplay between these forces has also been charted, their connection or lack of connection one with other, and the tensions between them which Locke either overlooked or failed to resolve in his evident struggle for coherence. In so far as the present attempt has been successful, a certain intelligible order has been revealed where the *Essay's* more recent readers (and some of its earlier ones) have been inclined to see only ambiguity and confusion. Some of the lacunae in the map will be filled in later chapters. For, as the variety and importance of the issues already discussed (however inadequately) might suggest, a full account of Locke's notion of an idea would be, in effect, an account of most of his general philosophy. Yet there is one virtual omission in particular which deserves to be mentioned at this point, and an attempt will soon be made to repair it.

That omission is of the topic, just touched on above, of Locke's intuitionism. 'Ideas', it should not be forgotten, are the 'materials' of reasoning and knowledge. They are what is before the mind when we understand or 'see' that something is so. Locke's whole conception of ideas as the logical constituents of conscious thoughts has the closest possible, mutually supportive link with his definition of knowledge as '*the perception of the connexion and agreement, or disagreement and repugnancy of any of our Ideas*'.¹⁴⁰ Some of the implications of that definition will be explored in later chapters. Nevertheless enough has been said about other aspects of the notion of an idea to justify a pause for recapitulation and reflection.

The first of the theoretical determinants of that notion presented above was the framework of traditional logic according to which ideas fulfil the role of parts of mental propositions, and which for Locke brought with it the notion of different categories of ideas, as also of names and of things conceived of and named. Next a several-sided dispute between traditional Aristotelian philosophy and some of its seventeenth-century rivals was summarily described, a dispute which ranged over epistemology and ontology but which was powered above all by the sense that a new view of nature forced a reappraisal of human knowledge and its

objects. Within the terms of this dispute it is possible to see how Locke's notion of an idea and his distinction between simple and complex ideas (in particular, complex ideas of substances) was tailored to fit his broadly mechanist, but at the same time sceptical or agnostic understanding of the world and of our relation to it. His principle that all our ideas come from experience served, within his system, both as a secure basis of our knowledge (since it is to its origin in experience that the absolute trustworthiness of a simple idea is due) and as an explanation of the limits of our knowledge (since simple ideas can be relied on to 'represent' or conform to the attributes of things only in so far as they are effects of those attributes). The essence or nature of things is left to speculation the terms of which are themselves tied to the ideas we receive through the senses. The best such speculation available to Locke he regarded (as we shall see) as demonstrably imperfect. Consonantly with this conception of an idea as essentially an appearance of something to the senses (or as an effect on the senses) Locke adopted the 'imagist' side of the dispute over the existence of a faculty of pure intellect.

Up to this point in the above account of Locke's concept of an idea no obvious inconsistency or incoherence had emerged, but a serious conflict was then identified between two lines of thought, one broadly Aristotelian or Scholastic, the other owing more to Epicurean and Stoic ideas which had only relatively recently been recovered and disseminated. First, much of Locke's language and thinking was found to reflect a duality in the notion of an idea recognized by Descartes and endemic to traditional ways of dealing with intentionality. As Descartes once put it, the term 'idea' is ambiguous, being employed *either* for representative modes of thought *or* for their objects as they are conceived of, as they exist objectively 'in the understanding'. Yet these were not taken to be really distinct, as Arnauld's convincing interpretation of Descartes stressed: an idea is something which is both representative mode of thought and intentional (or 'immediate') object at once, since to characterize either is to characterize the other. In particular, to give the intentional object of a thought is to give what Descartes called its 'form', i.e. to assert something about the thought and not about reality. Descartes and Arnauld further agreed that the intentionality of conscious thought is intrinsic and *sui generis*, an essential characteristic of immaterial substances:

this way of being objectively in the mind is so peculiar to the mind and to thought, in that it constitutes the specific nature of the mind, that one would look in vain for anything resembling it in whatever is not mind and thought.¹⁴¹

Digby was another philosopher who saw intentionality as tied to immateriality. Hobbes disagreed.

The second, Epicurean line of thought presented a more radical model for intentionality according to which a simple idea is the natural sign or

representative of its regular cause. As this model was developed in the *Essay* its tendency is to turn ideas into blank, 'immediately' perceptible sensory effects, i.e. mental effects each somehow possessing a sensible character in its own right. All sensible qualities of external things become in the first instance (and as far as we have sensitive knowledge of them) powers to produce such effects. That this whole approach is in tension with the traditional Scholastic conception of the 'immediate' objects of perception and thought was revealed in the consideration of arguments in the *Essay* which work, or at least seem to work, given the one model but not given the other: e.g. the analogy drawn between ideas of secondary qualities and pain, and the treatment of Molyneux's problem. The tension between the two models is epitomized in the uneasy ambivalence of the principle that ideas of primary qualities can be supposed to 'resemble' what they represent, as ideas of secondary qualities do not.

Reflections on the structure of thought

What, then, is the philosophical interest of all this for us today? It is tempting to answer, 'not much', on the grounds that the distance is too great between Locke's presuppositions and the insights which direct credible present-day approaches to intentionality and thought. Certainly, recent philosophy of mental representation tends to assign much more importance to language and behaviour, and much less reality or, at least, independence to conscious experience and thought than Locke did. To put it less neutrally, an exaggerated respect for the role of language and action in determining or constituting the content of thought, and a contempt, at least questionable, for consciousness now pervades philosophical theory. That Locke stood at the opposite extreme supplies one good reason for taking his arguments seriously.

One of the most radical modern criticisms of Locke's assumptions runs roughly as follows. If it is necessary, as he assumed, to attribute to each thought a structure related, element by element, to the logical structure of any sentence which accurately expresses that thought, then it is difficult to avoid the conclusion that 'thoughts' are mere abstractions. It is plausible, that is to say, that our talking of the 'thought' expressed by this or that utterance is just one way of abstracting from the total event or state of affairs of which the utterance was a part, in order to identify the speaker's achievement, *what was said* or *the saying*. For since the language spoken and the particular context in which the sentence was uttered will have made decisive contributions to what was said, there seems to be no role for an inner process of thought perfectly matching the sentence in its logical form. On the other hand, in Locke's favour, it is a blatant paradox to cut consciousness out of our account of meaningful speech altogether, or to reject the possibility of thought without language (which is not to deny that paradoxes can have considerable charm).

There appear, at any rate, to be two opposed models for the relation between language and conscious thought. On the one hand it has been a traditional assumption that consciousness is the primary bearer of intentional content, so that without appropriate states of consciousness uttered sentences are so much empty noise, spoken without understanding or meaning. On the other hand, it has seemed to an assortment of revered twentieth-century philosophers that language,

related as it is to reality by what are in effect social conventions or rules, has a life independent of the individual, so that inner states of consciousness are irrelevant to what is said and meant.¹⁴² In order to resolve this conflict and to begin to achieve a satisfying account of the relationship between thought and its expression, we need an explanation of how pre-linguistic or non-linguistic thought relates to expressed thought. We need to understand just how language can enter into and mould thought, bringing into existence whole categories of thoughts, such as mathematical and, in general, *a priori* thoughts, which would otherwise be impossible. That there is a problem here was not unrecognized in the seventeenth century, for Hobbes advanced an elaborate account of truth and universal reasoning directed, in effect, towards its solution. A shortcoming of his theory, however, as we shall see and as Locke evidently saw, is that the role ascribed by it to language was not in the end shown to be essential. So whereas Hobbes had defined truth in terms of names, Locke, while evidently taking something from Hobbes, saw no reason not to fall back on another, more traditional view: truth is the '*joining or separating of Signs, as the Things signified by them, do agree or disagree one with another*'.¹⁴³ There are, he held, two sorts of signs: ideas signify what lies in the world, while names signify ideas.

The notion of conscious thought prior to, or independent of, language is in ordinary life treated as unproblematic, but there are, or may seem to be, certain difficulties in it. If we ascribe a thought to an infant or to a dog, it is a thought 'that...'. In other words, the thought in question is identified precisely through our expressing it by means of a contained sentence with a certain logical form. The thought is in effect identified as that thought which is true if and only if the contained sentence is true. But it seems that the capacity of the sentence to be true can be understood, and needs to be understood, in terms of its logical form: crudely, it is true when what is referred to by the subject-term satisfies the predicate. That seems to be just the reason why Hobbes held that there is no truth or falsity without language. Yet (although Hobbes did not take this step) without truth or falsity there is no intentional content. Of course some thoughts, such as wishes and intentions, have content, but are neither true nor false. Yet wishes can be satisfied and intentions fulfilled, and a wish that *p* has the same content as a belief that *p*. Taking belief as, for present purposes, the simplest case of an intentional state, we can say that to give the intentional content of a belief is simply to state under what circumstances the belief would be true. That is the purpose of the contained sentence '*p*' in the ascription of a belief that *p*. So we seem to be faced with a dilemma: either the inner process of thought itself has logical form, as Locke and most of his predecessors supposed, and is a kind of inner speech; or without language there is no thought, because no truth or falsity, and so no content.

Here it should be said that, although present-day analytic philosophers of language and mind often see themselves as faced with a choice more or less correspondent to this dilemma, few would interpret that choice in the terms in which I have just presented it. That is because the possibility that it is inner

processes of thought which bestow meaning on outward utterances is now seldom if ever understood as the possibility that it is elements of consciousness, such as Lockean ideas, which fulfil that role. In the course of the twentieth century quite a new candidate for the status of the primary bearer of intentionality has emerged, i.e. action or behaviour. According to a recent and sophisticated form of behaviourism, so-called 'functionalism', the thought which lies behind an ordinary utterance, and which may determine, or help to determine, the meaning of that utterance, is a physiological state or process in the brain causally or functionally related to a certain input and a certain output. The input will ultimately be sensory input, and the output behaviour, verbal or non-verbal. It is this functional relationship, it is held, which makes it appropriate and possible to assign intentional content to inner processes, i.e. to conceive of any such process as the thought that *p*. To put the functionalist view very crudely, the thought that a pig is present is a physiological state or process of a kind liable to be caused by a present pig, and liable to give rise to behaviour appropriate to the presence of a pig, including (at least, in the case of language-users) such verbal behaviour as the utterance, 'A pig is present.'

It can accordingly be held that an utterance of the form, 'A pig is present', is indeed meaningful (or, at least, *meant* on some occasion) only if it is due to an appropriate inner process in the utterer, i.e. a physiological process, constituting an intention, which is systematically related to such a process or state as would constitute a belief that a pig is present. The precise form of this relation is subject to theoretical debate.

Given these proposals, it can be seen that there are, broadly speaking, two possible forms of functionalism with respect to language. First, it can be held that the kind of inner process which lies behind a meant utterance can be satisfactorily characterized, for the purpose of explaining what meaning is, entirely without reference to language. This is the view that the most fundamental explanation of intentionality will be by reference to the appropriateness of action to its circumstances.¹⁴⁴ A version of this approach holds that the inner process in question must itself be supposed to have a structure which is quasi-linguistic in character, a proposal which seems to echo seventeenth-century conceptions of a natural language of thought.¹⁴⁵

The second view is that such a reductive programme is impossible. It is impossible just because the articulation or form ascribed to any thought in ascribing it content can only be explained in linguistic terms. Consequently any functionalist account of the thought behind meant utterances would need to be given in terms of linguistic behaviour, or utterances of the very kind in question. That is to say, such an account would need to assume, and ultimately to refer to, just that context of conventional use which makes the sentence uttered generally meaningful and so *capable* of being meant when uttered. It follows that thoughts with determinate content can only be had by language-users.

Although our dilemma can be transposed along these lines into the terms of functionalism and modern theory, the possibility I want to explore here is better

introduced by the dilemma in its original form. Functionalism will be discussed at greater length below, and reasons will be given for rejecting it as a general philosophical explanation of intentionality and thought. For the present I want only to consider the claims of *conscious* experience and thought to be the primary bearers of intentionality, and even to be the ultimate source of logical form. The discussion of these claims is, of course, more directly relevant to a discussion of Locke's own theory, but it is also to the point that the claims of consciousness in this respect (as opposed to the claims of language or functional state) are those which are most rapidly passed over by present-day philosophers without serious argument. I want, in effect, if only in justice to Locke, to reopen a debate most philosophers seem to regard as closed. Ultimately the question hinges on whether, and in what way, consciousness and pre-linguistic cognitive capacities involving consciousness have to be mentioned, as well as public practice and convention, in an adequate account of what it is to understand an expression or learn its meaning. For perhaps the truth or falsity which presupposes logical form is dependent on a more primitive or fundamental kind of truth and falsity of which conscious states are in themselves capable. It is not so difficult, after all, to find in the relationship between conscious experience and its objects analogues for reference and the satisfaction of a predicate (and so for truth) without imputing fully determinate logical form to the experience.

Suppose that I see a face clearly and for what it is. The face is causing in me, in the normal and regular way, a visual experience or sensation which has a certain character, and in virtue of that character the experience 'fits' what is causing it. Significantly we can only describe the character of the experience in intentional terms. Most simply, the experience is a visual impression, or sensation, *of a face*, although there could be reasons for saying less: for example, that I am having an impression of *a pink patch*, and so forth. Very roughly, my visual impression, however described, has the face as its object (the pink patch *is* the face) just because (1) it is being caused by the face through the mechanism of sight and (2) it 'fits' or conforms to the face to a certain degree. This double relation constitutes a broad analogue to linguistic reference. The second condition is necessary because not every visual impression is necessarily 'of what causes it. After-images and experiences of mirages, and perhaps of rainbows, have external causes but present objects which do not exist.

The sensation need not, however, *perfectly* match its object. Indeed it can be quite misleading as to its object's real character, and yet still be 'of' that object. One way of perceiving things is to misperceive them, a possibility which supplies an analogue to the most straightforward case of falsity. Roughly, I misperceive, for example, a crumpled shirt on the back of a bedroom chair as a face when (1) the light from the shirt causes me to have a visual impression *of a face*, i.e. which a face would fit more or less and (2) it does so part by part, in such a way that there is a correspondence in spatial and certain other respects between the appearance of the face and the appearance which the shirt would have had, if it had had an unmisleading appearance. Note that we can describe

my sensation either as an impression of a face, or as a misleading impression of the shirt. And we can either *distinguish* the (existing) shirt from the (non-existent) face, or we can *identify* them (and attribute existence to them both) in saying, 'The face was just a shirt'. Here we may be reminded of Cartesian problems, discussed above, as to the relationship between the sun as it exists in the mind and the sun as it exists in the sky. Note, too, that while it is possible to misperceive a crumpled shirt as a face, the second condition of misperception means that it is not possible to misperceive a flash of light as an elephant, or a straight line as a pentagon, even if the perceiver is in an abnormal psychological or physiological state such that the presence of the light or the line before his eyes causes him to have the visual impression of an elephant or a pentagon. Such an experience would be an hallucination with an external cause but with no real object, whereas the misperception of a shirt as a face is better described as a visual illusion.

These considerations may not in themselves advance the argument very far. Even if certain causal *cum* intentional relations between sensation and its objects do constitute forms of reference, truth and falsity more primitive than linguistic reference, truth and falsity, a certain obstacle must be surmounted before we can understand how far the former generate the latter, or even underlie their possibility. Sentences express thoughts, not sensations. The intentionality of sensory experience may seem too distant from full-blooded prepositional thought for it to do the necessary work. On the other hand, of course, we can only refer to what we have some knowledge of, and we have knowledge of objects through the senses.

Perhaps the first or most obvious task facing a philosopher concerned with the relationship between sensation and thought is to explain memory, or the possibility of thinking about what is no longer present to the senses. It is of course one of the main attractions of what I have called 'imagism' that it offers some account of that relationship. According to imagism, something essentially *like* a sense-experience or sensation arises in memory or is formed in imagination. An image represents in the same *sort* of way as sensation represents. If it arises in the memory, then, like sensation, it may either be faithful, or relatively unfaithful to the reality from which it derives. If it constitutes mere imagination, then it may either have no connection with any real situation, or it may be connected to some situation more circuitously than in the case of sense-perception or memory (as when I imagine, accurately or inaccurately, what lies on the other side of a wall).

The imagist story will have to be evaluated. To that end it must be considered whether it is necessary to include any other type of mental representation than the primitive intentionality involved in sense-experience in an account of pre-linguistic conscious thought in general. As for *a priori* thoughts, it will be argued below that imagist explanations such as Locke's, however ingenious, are demonstrably inadequate. As it has long been recognized, language has here an essential role to play. Mathematics, for example, is inconceivable without

symbolism, as Locke himself understood in the case of arithmetic. But it is difficult to arrive at a satisfying characterization of that role. In effect, the question of the sort of representation or intentionality involved in *a priori* thought is identical with the question of the nature of the understanding of language itself, as employed by oneself or others. The inadequacy of an imagist account of the one implies the inadequacy of an imagist account of the other. But if an imagist account of meaning and understanding is inadequate, it follows that either there is no immediate relationship between the intentionality of consciousness and the intentionality of speech, or else the understanding of speech involves some form of representation at the level of consciousness other than sensory or quasi-sensory representation. All these questions (including the question as to what 'images' may be) will be considered below in the course of a more detailed critical response to Locke's own theory of perception and thought.

A more fundamental problem, however, requiring resolution if the logical form of linguistic utterances is to be perspicuously founded on the intentionality of consciousness, is that of identifying some articulation or structure in the content of consciousness capable of generating the articulation and structure of language. It is not enough to say that sensations may 'fit' or fail to fit what causes them. Indeed my comparison of the 'reference' of sensation and its 'fitting' its object with linguistic reference and the satisfaction of predicates is open to the criticism that I have simply assumed such articulation. To talk of the visual impression of a *shirt* or a *face* is already to assume that the world is presented to sensation as divided up into distinct objects resembling or differing from one another in determinate ways. The crucial question remains as to just how far, in giving an account or description of the content of a sensation or experience, we impose or project articulate form and determinacy onto that content.

Locke's answer is that we do not do so at all. He assumed that there are elements of any experience, simple ideas, which are no less determinate than the elements of any true account of the experience. Indeed they *cannot* be less determinate, since the terms employed in such an account get their meaning precisely by association with such elements of experience, alone or in combination. Moreover, corresponding to the subject-predicate form of the contained sentence in, for example, 'He perceives that the gold is yellow', the experience ascribed has, according to Locke, the form of a perception of the coexistence of qualities in a substance.¹⁴⁶ We can in general, Locke believed, look into our thoughts and analyse and categorize both the ideas which we find there and the logical form of the mental affirmations in which those ideas are employed. The way in which, according to him, form or articulation is imposed by us where it has no real existence is not in the description of experience, but in the experience of reality. He believed, as we have seen, that sensible 'accidents' have no separate or distinct existence in reality. Sense-experience does, however, present some genuine or 'given' elements of reality, namely substances, although it is doomed to do so under the form of a spurious complexity. As Digby put his similar view, 'what is but one entire thing in it self, seemeth to be many things in

my understanding'. That the mind-dependent many things belong to the independently unitary 'entire' thing was apparently for Locke something grasped by the senses themselves, in the perception of 'coexistence'.

Again it may be helpful to oppose Locke's view of these matters to the currently popular view, very different from his, that it is language that imposes form or articulation both on experience and on reality. That is to say, no objects are 'in themselves' distinct and individual, nor are they perceived as such independently of language or of 'concepts' embodied in language. To suppose that reality or experience comprises certain categories of entities (substances, events, qualities, sense-data or whatever) is, on any view of this kind, simply to see it through the medium of a particular language, 'conceptual scheme', 'theory', 'language-game', 'practice', 'ideology' or the like. There is no such thing as absolute ontology. All logical form, of whatever kind, belongs in the first instance to language and is imposed, through the medium of language, on what lacks it in itself. There are no 'given' objects. This doctrine, which might be called 'linguistic idealism' or 'linguistic conceptualism', will be subjected to some oblique criticism in the present volume, and to a more sustained and frontal attack in the sequel.

Part II

Knowledge and belief

10

Introduction to Part II

In an age devoted to systems of classification, seventeenth-century epistemology was no exception. Yet the various classifications of the cognitive faculties and of the species of knowledge and opinion were constructed with an eye on ancient paradigms as well as current controversies. In order to understand the nature and point of Locke's own enterprise it is necessary to be aware of at least a part of that background.

Parmenides was perhaps the first Greek philosopher to draw a technical distinction between opinion and knowledge, together with corresponding distinctions between appearance and reality, and becoming and being. Plato interpreted the distinction in such a way that, while opinion (*doxa*) relates to the changeable world of sense-experience, knowledge (*episteme*) is of the universal principles of explanatory science, having as its object immutable transcendent forms. Apprehension of the universal can be prompted by the deliverances of sense, but is itself the operation of a higher faculty. Plato sometimes suggested that the forms were directly apprehended before birth, so that in this life we grasp universal truths by recalling them. That thought introduced the notion of innate knowledge to European philosophy.

Aristotle retained the distinction between universal science, concerned with the forms of things and achieved by the intellect, and sensory or experiential knowledge. Nevertheless the former is founded on the latter, and is neither innate nor concerned with transcendent universals. In a famous passage he charted a four-stage progress to science: first, the perception of particulars; second, the memory of single particulars; third, 'experience' formed from the memory of many particulars and constituting the apprehension of what is in common to them; and, fourth, the apprehension of universal explanatory principles, the definitions used in scientific 'demonstration'.¹ Yet although this account gives a more important role to the senses than Plato had done, indeed an essential one, we are nevertheless supposed to grasp something universal in itself, even if not transcendent, and to achieve an intellectual comprehension of principles which is qualitatively different from fallible opinion.

The Epicureans and, more expressly, the Stoics rejected both transcendent and immanent universals: for the latter, the principles of universal knowledge are

simply the generic concepts or 'preconceptions' impressed on the soul by 'experience' or 'many memories of a similar kind'. Consequently, for all the echoes of Aristotle's model, the Stoics did not share his sense that, as we progress from particular perceptions to science, there is a qualitative shift in the nature of our apprehension and of the faculty employed. Other features of their epistemology owed more to their confrontation with the Sceptics and to their attempts to characterize a 'criterion' or cognitive state which is infallible, the operation of a natural faculty to discriminate truth from falsehood. One of the tools they employed was the notion of a faculty of assent. Briefly, some of our sensory impressions are 'cognitive' or 'cataleptic' impressions which 'grasp' their objects, as when we see something clearly and distinctly before us. At least when there is no countervailing reason, such impressions naturally and inevitably command assent. But it is possible, on the one hand to give rash and precipitate assent to incognitive impressions, constituting opinion, and on the other to advance beyond cognitive impressions to systematic scientific knowledge which is absolutely secure. It is also possible, of course, to withhold assent, as the Sceptics recommended in every case. The Stoic wise man only assents when he has scientific knowledge, otherwise withholding assent. The terms 'opinion' and even 'ignorance' were sometimes extended to all assent less than science, even to natural assent to cognitive impressions given in circumstances in which their status is not put beyond doubt. At the same time, cognitive impressions are the foundation of all knowledge, and sense-perception in ascertainably normal conditions, on the part of someone with the appropriate discriminatory skill, is the ultimate arbiter of truth.²

We can perhaps trace the birth of the epistemology which now seems characteristic of the seventeenth century, in so far as it is at all possible or helpful to impose such beginnings on the history of philosophy, to the vigorous revival of the ideas of Hellenistic philosophy: first, in the sceptical writings of Montaigne and his followers towards the end of the sixteenth century, and then in the reactions to them. An early and effective response to the view that nothing is known was that of Mersenne, who argued that not only are appearances certain, in that, if the sunlight feels warm, it is certain that it does so, but that reason can move beyond appearances by considering the circumstances and by devising appropriate checks, so as to arrive at safe perceptual judgements as to what actually *is* the case. Rational judgements have their own light, so that it is inappropriate to ask, as the sceptic does, for a justification of trust in reason. No one can really doubt such maxims as that it is impossible for the same thing both to be and not to be, or that the whole is greater than the part, or that we should shun evil and pursue good. Neither sense nor reason reveals the essences of natural things, but reason can at least sometimes move from effects known through the senses to arrive, by means of evident principles, at certain judgements about what is beyond the limits of perception. When that is impossible, we can often form probable judgements. Abstract truths, however, as in pure mathematics, can possess an absolute, self-illuminating certainty which

assures us that in reason we have a faculty capable of science, the highest degree of knowledge. Although Mersenne did not claim knowledge of natural essences, he was an advocate of a mechanistic and mathematical approach to nature, and he included among self-evident maxims the proposition that the world comprises bodies in motion. The tendency of his epistemology was to raise reason above the senses both as the judge of their deliverances and as an independent source of science.³ This tendency was developed by his friend Descartes.

Descartes' epistemology, already discussed in part above, seems in many respects to have been shaped by Stoic theory, starting with their assumption or conviction that nature has provided us with a faculty for distinguishing truth from falsehood. Assent is voluntary, and precipitate assent to anything less than 'clear and distinct' ideas is error, including even our natural inclination to believe the normal deliverances of the senses when that inclination has not been endorsed by reason. It is entirely unsurprising that he chose to expound his theory in the context of an internalized dispute with a sceptic. Where he sharply differed from the Stoics was in his assigning clear and distinct ideas to a source other than the senses, and in his associated conception of essences and eternal truths, the objects of science. Although he held with the Stoics that 'all universals are simply modes of thinking',⁴ he adopted, *mutatis mutandis*, a form of the extraordinary Scholastic harmonization of Platonic, Aristotelian and conceptualist theories of universals according to which the forms of particular things are themselves particular, while universal essences are ideas in God's mind, the archetypes of his creatures. That model itself involved other conceptions of science than those considered above: we possess knowledge in so far as God reveals His mind to us, or has formed our minds in the image of His own.

A formal classification of knowledge and opinion on Cartesian lines was offered in the Port Royal *Logic*, if modified by its authors' orthodox conception of 'judgement' or affirmation (discussed in [chapter 2](#), above) as an act of mental proposition-making, the combination of concepts. The *Logic* identified two main sorts of affirmation or denial, 'conviction' and 'opinion'. The former may be intuitive and immediate knowledge, as of a self-evident maxim or first principle. Otherwise it may be motivated by reasons, good or bad. If by good reasons, it is 'science' or understanding. If by bad reasons, it is error or rash judgement. Yet a third sort of 'conviction', motivated by authority, human or divine, is 'faith'. 'Opinion' arises when judgement is motivated by reasons which are less than convincing: it is an 'acquiescence of the mind accompanied by doubt'.⁵ One disadvantage of this scheme is that it fails to explain how reasons for a judgement could be assessed as good or bad if no proposition is before the mind prior to judgement. Nor does it explain how a belief's being motivated by reasons is related to its being motivated by something irrational, such as passion, wishful thinking or interest. And there seems to be no room in the scheme for someone's finding conclusive reasons less than fully convincing.

Gassendi was another philosopher who saw affirmation as an act of proposition-making, but, as we should expect, he proposed a classification radically different from the *Logic*'s. He held that judgement arises when the mind joins ideas in affirmation as they fit one another (*mutuo congruent*) or when it separates them in negation as they fail to fit. An orthodox distinction was drawn between demonstrative syllogisms, of which the premises are necessary and evident, and probable syllogisms, of which the premises are contingent and likely. The former issue in 'science', the latter in 'opinion', a kind of provisional assent. Where Gassendi most sharply diverged from Mersenne and Descartes was in emphasizing the Epicurean view that the greatest 'evidence' is the evidence of sense, upon which, indeed, all other evidence depends.⁶ A proposition has 'verisimilitude' when, as a result of past experience and, in some cases, the testimony of others, it 'approaches evidence more nearly than obscurity'.⁷ A distinction, he continued, is commonly drawn between a *a priori* demonstration, of which the premises are both universal, and a *a posteriori* demonstration, of which one premise is particular. These terms he condemned as quite misleading, since they imply that the most general propositions are the first known, whereas particulars are known first and generals afterwards. All the evidence and certainty of a general proposition derives from that of the particular and less general propositions which come under it: the evidence of 'Every whole is greater than its part' depends on the evidence of such subordinate propositions as that the whole sky is greater than one star.⁸ This argument had its influence on Locke, as we shall see, but Locke did not embrace the extreme empiricist view of necessity and universal knowledge to which Gassendi appears to have been committed. Finally, like Arnauld, Gassendi saw 'faith' as a separate type of judgement. Yet he made the point that we can and often do have faith and opinion about the same things, since the acceptance of (at any rate, human) authority will involve an opinion as to the speaker's veracity in saying what he does.⁹

Another approach hinged on the role of language in thought: for Hobbes, affirmation is a linguistic act and all propositions are linguistic entities. Names are signs of conceptions or ideas, but sentences do not mirror an internal and combinatory act independent of language. There is no such generic mental act to be designated 'affirmation' or 'judgement', of which intuitive knowledge, opinion and so forth are the species. On Hobbes' scheme, knowledge of fact, a certain kind of 'judgement' and a sort of error are possible without language, but affirmation, truth and falsity, universality, science and reasoning are made possible only by the introduction of names. A proposition is formed only when 'of two appellations, by the help of this little verb *is*, or something equivalent, we make an affirmation or negation'.¹⁰ The proposition is true when that which is named by the subject is named by the predicate. Its truth is 'evident' when we actually have an idea or image of the thing named, and so can recognize that it is the same thing properly named in each case.¹¹ We shall consider some further details of the theory, and the nub of the issue between Locke and Hobbes, when

we come to examine the relationship between Hobbes' nominalism and Locke's doctrine of abstraction. What is relevant to the present topic is Hobbes' classification of the cognitive states or acts.

Hobbes drew two main, intersecting lines of distinction: between knowledge and what is less than knowledge; and between states which do, and states which do not presuppose language. The 'knowledge' which is possible without language is 'knowledge of fact'. This is 'sense and memory', 'the knowledge required in a witness'. There may also be in us, or in animals, a 'presumption' about the past or future, which is 'opinion'. 'Opinions' are a sort of inclinations naturally arising in consciousness as a result of past experience of circumstances like the present (Hobbes was without doubt an important influence on Hume). A chain of alternating opinions is called 'doubt', and its final result is 'judgement'. To judge well is to be 'prudent'. Hobbes found room for a broad analogy between judgement and voluntary action, but on his particular form of psychological determinism or naturalism it consists simply in the analogy between the presumptions or opinions which issue in judgement and the 'appetites' which issue in 'will' ('the last appetite in deliberation'), and so in action. Hobbes' language-dependent cognitive states include the second form of knowledge, namely 'science' or 'the knowledge of the consequence of one affirmation to another'. Science arises when a truth becomes evident through syllogistic reasoning from definitions. If there is an error in the reasoning, acceptance of the conclusion constitutes a second sort of mere 'opinion'. Finally, to accept what another says, i.e. to draw conclusions from the speech of another, is to have 'faith' or 'belief'.¹²

All these systems, and others, had some effect on Locke's own classification, which was possibly the most complex of all. His central distinction is between knowledge and opinion, which he also called 'belief' and 'judgement', like Hobbes separating 'judgement' from 'knowledge'.¹³ 'Perception' is not, as by Arnauld, restricted to the perception of ideas taken individually, for knowledge is itself the perception of 'the connexion and agreement, or disagreement and repugnancy of any of our Ideas'.¹⁴ It is the perception which itself constitutes the mental act of affirmation. The mind 'judges' something to be so when two ideas not perceived to be evidently connected are 'presumed' (the same word as Hobbes had used) to be connected on the basis of extrinsic 'arguments or proofs' which make the proposition 'pass or be received for true'.¹⁵ Someone who follows a geometrical demonstration *knows* its conclusion to be true, while the person who accepts the conclusion from a reputable mathematician only believes it. The same distinction can be drawn with respect to matters of fact: if I see someone walking on ice I know that he does so, whereas if I am told of his doing so I may believe it or not, depending on my past experience and the circumstances of the telling.¹⁶ Thus although Locke's categories of knowledge and belief are mutually exclusive, the distinction between them had for him a status significantly different from its status for much previous theory. For him, the subject-matter of knowledge and belief is in general the same: not only is it

possible (with a certain exception) to believe what can otherwise be known, but it is in principle possible to know whatever can be truly believed.

Locke divided knowledge according to several different, if interrelated principles. First, he distinguished between four types of proposition known.¹⁷ Second, he distinguished between ‘actual’ knowledge and two types of ‘habitual’ knowledge.¹⁸ Third, he distinguished three ‘degrees’ of knowledge.¹⁹ Fourth, he divided up knowledge according to the category of its objects, as it concerns simple ideas, substances, modes or relations.²⁰ Finally, there is his special account of the distinction between particular and universal knowledge. Each division was drawn up with the others more or less in mind and each reflects a good deal of theory, for Locke was fighting, as ever, on several fronts.

Some of these divisions of knowledge will be discussed in chapters 11 and 12. The distinction between particular and universal knowledge, however, has perhaps been sufficiently introduced, and will receive extensive consideration in Part IV. The categorization of knowledge according to subject-matter will not be separately discussed here, but will be a main topic of Volume II. The most significant distinction by subject-matter is between knowledge of substances on the one hand, and knowledge of modes and relations on the other, a distinction which constitutes Locke’s explanation of the distinction between the natural and the *a priori* sciences. Chapters 13 and 14 deal with Locke’s conception of belief and assent, with particular respect to his explanation of error and his account of the grounds of probability. His attitude to revelation and faith will be an important topic. In chapter 15 an attempt will be made to relate Locke’s approach to knowledge to what may seem the very different, even incommensurable perspective of many present-day writers. The idea will be introduced (to be explored further in Parts III and IV) that the shift in the theory of knowledge has involved not inconsiderable loss, as well as some gain. Chapter 16 will supply a similar, if much shorter discussion of the philosophical significance of Locke’s largely neglected treatment of belief and error, and will explore the source of the tensions in his account. The difficulties he met with will be related to the issue of the virtues and limitations of so-called ‘naturalistic’ approaches to belief, an issue broached in the discussion of knowledge.

The degrees of knowledge and the role of method

The exception to the principle that whatever can be known can be believed is the highest degree of knowledge, intuitive knowledge. What can be known intuitively cannot be believed because in intuition the mind immediately perceives a relation between ideas ‘as the Eye doth light’. There is no room for ‘judgement’ that black is not white, or that three is one more than two, or that I myself exist. To distinguish and compare the ideas concerned is immediately to perceive, and so to ‘know’, their relation. In this kind of knowledge there is ‘no room for Hesitation, Doubt, or Examination’, or for error.²¹

Next comes ‘demonstrative’ or ‘rational’ knowledge, where a chain of intuitively perceived steps is necessary for us to perceive the relation between two ideas. The chain is constructed by means of ‘intermediate ideas’. Locke gave the standard example of the Euclidean demonstration, employing diagrams, that the angles of a triangle are equal to two right angles.²² The ‘intermediate ideas’ are presumably the further lines and angles constructed in the diagram for the purpose of the proof. Their purpose is like that of a measuring rod: ‘As a Man, by a Yard, finds two Houses to be of the same length, which could not be brought together...by juxta-position.’²³ Locke did not think of the need for such steps as variable from person to person, which would suggest that the need stems from the relation itself. That would be consistent with his view that propositions known intuitively could not conceivably require a proof for anyone who understood them; but once in the *Essay*, unfortunately, he indulged in the speculation that demonstrative knowledge for man might be intuitive for angels.²⁴ For us, at any rate, prior to demonstration there is room for doubt, ignorance or belief about the conclusion, and so room for error. Even after demonstration, it seems, the sheer complexity of the proof can lead to ‘a great abatement of that evident lustre and full assurance’ enjoyed by intuition. In the course of demonstration, moreover, memory is commonly necessary. Because in long proofs we may think we remember previously perceiving connections where there are none, ‘this is more imperfect than intuitive Knowledge, and Men embrace often Falsehoods for Demonstrations’.²⁵ Nevertheless, demonstrative knowledge is knowledge because it involves only what is known by intuition, even if remembered intuition.

Because Locke accorded the greatest certainty to immediate intuition, his theory of knowledge has been regarded as essentially Cartesian.²⁶ No doubt his conception of certainty owed something to Descartes, for in *Draft A* he agreed that 'I exist' is the most certain of all propositions; while his point that the conclusion of demonstrative reasoning rests on memory is a Cartesian theme. So too the word 'intuition' seems to echo Descartes, who claimed novelty for his use of the term '*intuitus*'.²⁷ Yet these connections should not be exaggerated. A distinction between the 'intellection' by which we know axioms or maxims and the 'ratiocination' by which we know the conclusions drawn from them was entirely traditional. Even something like the speculation that angelic knowledge is all intuitive had its place in the logic books.²⁸ Moreover, like Gassendi, Hobbes and ancient empiricists, Locke extended the terms 'evident' and 'certain' to sensory knowledge. Intuition is not for him an act of pure intellect, but is tied essentially to what Descartes described as 'the fluctuating testimony of the senses' and 'the blundering constructions of the imagination'.²⁹ That could even have been the reason why Locke (like Gassendi) preferred to speak of 'intuition' rather than 'intellection'.

The doctrine of 'intermediate ideas' which are employed in demonstration, and Locke's thoughts about how to hit upon such ideas, also grew from deeper and more traditional roots than certain passages in the *Regulae* which, indeed, he may never have seen. As almost any logic of the time would explain, the method for solving a 'question' involves seeking an 'argument' or 'medium' or 'middle term' with which to construct a syllogism. The middle term *animal*, for example, links the 'extremes' involved in the question whether man is mortal. 'Invention' was the part of logic concerned with the finding of arguments, and a more or less elaborate list of 'topics' or '*loci*', originating in Aristotle's work of that name, gave the places to look. In our example, the argument or medium *animal* is simply classifiable as the *genus*: it derives from the topic, *genus*. In a passage from Virgil used by Pierre de la Ramée and discussed in the Port Royal *Logic*, a speaker says (to put it more briefly, if less elegantly than the original), 'Kill me, not him: I led him to do it'. This was classified by Ramée as an argument from the efficient cause, and for Arnauld and Nicole it illustrates the uselessness of the 'topics' for discovery. It would be absurd to suppose that the speaker might have worked studiously through a list of topics in order to hit upon such a reason for requesting to be killed. Reasons come naturally to us as the product of common sense and knowledge of the subject: logicians' classification comes later. The traditional theory of 'invention', they conclude, is no real contribution to method.³⁰

Similar criticism was brought against syllogism in general, although, as we might expect, rather less vigorously by the authors of a popular logic than by Descartes himself. For Descartes the forms of syllogism are frames on which our thought can, if we wish it, be stretched, but few ordinarily think like that and nothing of practical value is achieved by trying to do so.³¹ The method of analysis and synthesis advanced in the *Rules* was indeed a method for solving

'questions', but was regarded as entirely independent of syllogism. Descartes gave as an example the question as to how a 'Tantalus cup' is made. This vessel contained an image of a man bent over to drink, and was so constructed as to empty completely as soon as the water rose to the image's lips. Analysis of what happens in such a cup reveals its general character, enabling the physicist to see that the image is irrelevant, and that only levels count. He can then construct an explanation, the presence of a syphon in the cup, from which the effect would necessarily follow. Analysis breaks down the subject-matter to its general principles, synthesis starts from these principles to explain the particular consequence.³² In a later work Descartes confessed that his physical principles 'are so simple and general, that I notice hardly any particular effect of which I do not know at once that it can be deduced from them in many different ways.' The solution is to seek further experiments 'whose outcomes will vary according to which of these ways provides the correct explanation'.³³ Yet such initial uncertainty, supposedly capable of resolution by experiment, was not held to extend to the highest principles, the fundamental laws of motion which can be revealed by intellectual analysis. What is variable in these rival hypotheses is the specific modification or structuring of matter, not its general nature.

The conception of method as analysis and synthesis was dominant throughout most of the seventeenth century. It is present in Bacon, although Bacon assigned a wider and more continuous role to systematic observation and experiment than Descartes allowed for. Yet not all philosophers, as we have already seen, were prepared to agree that what is most general is most certain, or that the more general bestows certainty on the less general. For Gassendi, all general principles are intrinsically less certain than the particular facts to be explained. For others there was sometimes another point of disagreement with Bacon and Descartes. The pyramid or hierarchy of knowledge, with a way up (analysis) and a way down (synthesis), is a model which has a history going back at least to Plato and which permeates Aristotelian doctrine. Consequently it is not surprising that Hobbes found it compatible with a respect for syllogism as the way to reason with clarity and precision.

Locke's position on these issues is not wholly easy to plot. There is no doubt, however, that he sided with Descartes and Bacon on one question, for much of a long chapter was devoted to the argument that syllogism is neither the 'proper instrument' of reason nor 'the usefulest way of exercising this Faculty'.³⁴ He 'readily' agreed that all valid reasoning may be expressed in the forms of syllogism, but held that 'those artificial and cumbersome Fetters' which 'clog and hinder the Mind' do not represent the quickest and clearest movement of thought. The clearest mathematical proof is palpably not syllogistic. Locke's rival model is that of a sequence or 'chain' of ideas or concepts between each pair of which the mind perceives a connection. For example,

the Mind seeing the connection there is between the *Idea of Men's punishment in the other World*, and the *Idea of God punishing*, between

*God punishing, and the Justice of the Punishment; between Justice of Punishment and Guilt, between Guilt and a Power to do otherwise, between a Power to do otherwise and Freedom, and between Freedom and self-determination, sees the connexion between Men, and self-determination.*³⁵

To render such reasoning syllogistically is to upset its natural order and to make the simple complex. Even those expert enough to do so will first have reasoned naturally. Locke probably arrived at this 'chain of ideas' model by reflecting on Hobbes and Gassendi, or indeed on the traditional theory of the syllogism, quite as much as on Descartes. For example, Gassendi's discussion included a criticism of the conventional ordering of the premises. Placing the major premise first of a syllogism in barbara confuses the natural order of thought: 'Man is an animal, animals are mortal, therefore man is mortal', is a preferable order, since the middle term comes in the middle.³⁶ In the *Essay* this modest plea was repeated as it stands, but in another guise it seems also to have been at work in the earlier argument that thought does in fact have a natural order.³⁷

Descartes' chief, if not sole, objection to syllogism was that it is useless for finding things out, but Locke's 'chain of ideas' model had a different point. 'Invention' had been standardly distinguished in Aristotelian and Ramist logic from the 'ordering' of arguments for the purpose of making a proof clear and judgement easy. Locke correspondingly distinguished 'sagacity' and 'illation' as aspects of rationality, arguing that syllogism is an effective tool of neither.³⁸ The 'chain of ideas' is a rival suggestion with regard to ordering or 'illation'. His attitude to 'invention', the chief concern of Cartesian method, was rather more negative. As he turned to it in the chapter 'Of Reason', he remarked that he did 'not pretend to have found or discovered here any of those right helps of Art' spoken of by earlier writers.³⁹ He can only encourage better men than he to search for 'new and undiscovered Ways to the Advancement of Knowledge'. Dogmatic faith in the potential of an analytic method is to be rejected, for the analysis of the terms of a question can never take us beyond the limited ideas we have.⁴⁰ The same rejection is implicit in an earlier chapter on method, 'Of Maxims'. Even the move to such general and abstract truths as are accessible to us is denigrated in a polemic against the common belief that 'maxims' or 'axioms' can serve as the foundation of science. Less general propositions logically derivable from received maxims are, if the maxims are true, just as evident in themselves and 'easier and earlier apprehended'.⁴¹

the Child, when a part of his Apple is taken away, knows it better in that particular Instance, than by this General Proposition, *The Whole is equal to all its Parts*; and...if one of these have need to be confirmed to him by the other, the general has more need to be let into his Mind by the particular, than the particular by the general.⁴²

Locke conceded to maxims only another, less ambitious use. Gassendi, following in a tradition, had distinguished three sorts of method: not only *methodus inventionis* and *methodus iudicii* but *methodus doctrinae*. Locke similarly argued that those who see the sciences as founded on maxims and general *praecognita* fail to ‘distinguish between the Method of raising any Science, and that of teaching it to others as far as it is advanced’.⁴³

The really pernicious error, however, he took to arise in physics. Any general principle or ‘axiom’ about the natural world reached purely by analytic abstraction, e.g. a definition of the essence of *man* or of *matter*, is necessarily arbitrary and useless. It is a hypothesis ‘laid down at pleasure’ and tailored to fit preconceived conclusions.⁴⁴ The objection is elaborated in yet another chapter specifically on method, ‘Of the Improvement of our Knowledge’. When the objects of our inquiry are substances, ‘we advance not...by contemplating our Ideas, and considering their relations and Correspondences.... Experience here must teach me, what Reason cannot.’ No mere method can convert descriptive natural ‘history’ into explanatory ‘science’. Hypotheses may ‘direct us to new discoveries’, but even then they are at best probable. Dogmatic philosophers have merely been deceived by ‘the Name of Principles’ into receiving ‘that for an unquestionable Truth, which is really, at best, but a very doubtful conjecture’.⁴⁵

In the same chapter of the *Essay*, however, the division drawn between the *a priori* and the empirical sciences seems to have been accompanied by rather more optimism about the possibilities of a method of ‘invention’ in the former. Locke did nod in the direction of Cartesian rules to the extent of finding in the procedure of mathematics a certain, if vague model for the ‘sagacious and methodical application of our thoughts’ starting ‘from very plain and easy beginnings’, a method which might conceivably be transplanted to the *a priori* science of ethics. Yet the only instrument of invention definitely identified is algebra, which it is difficult to see significantly operative in ethics.⁴⁶ It seems in general that for Locke ‘the Art of finding out those Intermediate Ideas’ was an improvement of natural sagacity which comes with practice, rather than something to be set out in rules. Certainly he devoted considerably more space and emphasis in the *Essay* to another aspect of method which had also received much discussion by others, but which is less relevant to our present concerns. That is the Baconian (or much broader than Baconian) programme for the reform of scientific language, the achievement of precise, settled and generally agreed meanings for our terms, based where appropriate on careful observation.⁴⁷

The third and last degree of knowledge is sensitive knowledge ‘of the existence of particular external Objects’, where in actual sense-experience ‘we are provided with an Evidence, that puts us past doubting’.⁴⁸ The inclusion of this type of knowledge, and the manner of its inclusion, is central to Locke’s philosophical position. First, his counting it as knowledge at all placed him with Gassendi and Hobbes, against Descartes. As ‘knowledge’ it possesses intrinsic ‘evidence’ and ‘certainty’. Yet, if there are degrees of ‘evidence’, it might seem

that 'evidence' cannot sharply distinguish knowledge from opinion just because degrees of evidence must be supposed to be, or to be continuous with, degrees of probability. If absolute certainty lies at one end of the spectrum, then anything less, it might be supposed, must lie where Gassendi placed probability, between the extremes of absolute evidence and absolute 'obscurity'.⁴⁹ Locke himself may appear to have implied as much by his remark that sensitive knowledge, 'going beyond bare probability, and yet not reaching perfectly to either of the foregoing degrees of certainty, passes under the name of Knowledge'.⁵⁰ Moreover, we are told that the highest probabilities rise 'so near to Certainty that they Govern our thoughts as absolutely, and influence all our Actions as fully, as the most evident demonstration'.⁵¹ Probability 'upon such grounds carries so much evidence with it, that it naturally determines the Judgement, and leaves us as little liberty to believe or disbelieve, as a Demonstration does, whether we will know or be ignorant'.⁵² Yet this atypical use of the word 'evidence' for the force of mere probability extended Locke's standard use to the point of ambiguity. As the careful wording of the sentence in which it occurs should remind us, to attribute to Locke the notion of a continuum between knowledge and belief, with 'sensitive knowledge' squeaking into a place above the line, would be to fail to take into account the absolute nature of his distinction between them, and so to miss the whole point of his classing sensitive knowledge with the former.⁵³

There is no doubt a problem as to how there could be such an absolute distinction and, at the same time, degrees of 'evidence'; a problem to be discussed below. But Locke's view, at any rate, was that the 'evidence' of knowledge is somehow intrinsic to that relation between two ideas which is 'perceived'. 'Knowledge' can be dependent on no such extraneous and contingent considerations as conformity with past experience or the trustworthiness of a witness. His thesis about perceptual or 'sensitive' knowledge, then, was that it is immediate and non-inferential: the deliverances of the senses have primitive authority underived from reasoning or the judgement of probabilities. Hume, in his discussion of our belief in external objects, distinguished those philosophers who hold that the senses themselves present their objects as independent, from those who assume that reasoning produces the belief.⁵⁴ Locke, like Hobbes and Gassendi, was a philosopher of the first kind, not the second. Sense is a natural cognitive faculty side by side with reason, if not prior to it. It is not, as Descartes held, the mere provider of material which reason must interpret and endorse if we are to have knowledge.

On the other hand, Locke deliberately (as we must suppose) rejected Gassendi's doctrine that the deliverances of sense possess the *highest* degree of evidence. For Gassendi, as we have seen, the evidence even of such principles as that the whole is greater than the part derives from the evidence of sense. Locke, however, was evidently sufficiently impressed by the claims of the intellectualist tradition on behalf of *a priori* knowledge to want to mark off a middle way between Gassendi and those such as Mersenne. Knowledge that the whole is greater than the part is explained as *a priori* in the sense that its 'materials' may

be ideas of imagination just as well as ideas of sense. The perception of a relation between two ideas, even when these happen to be ideas of sense deriving, for example, from actual geometrical diagrams, is supposed different from, and in a certain way less questionable than, the perception that these ideas *are* deriving from reality here and now— that a triangle and an angle equal to two right angles actually exist before me. Particular existential propositions tie the framework of ideas to reality. It is his adoption of this model which explains why Locke defined ‘sensitive knowledge’ as knowledge of particular *existence*, for on his account the senses could be the faculties involved in intuition or demonstration, but only in knowledge of existence is their role essential.⁵⁵

The part assigned by Locke to the senses in the acquisition of knowledge will be considered in more detail in [Part III](#), below. Given the category of ‘sensitive knowledge’, it may seem a remarkable assumption that knowledge, when we have it, is the product of faculties in themselves not only authoritative, but infallible. The notion of degrees of knowledge does not imply that the ‘perception’ involved in demonstrative and sensitive knowledge is fallible. What Locke supposed to vary in degree, since it is not probability, seems to be security from a certain sort of error: not the error of perceiving what is false, for that is impossible, but the error of taking ourselves to perceive (or to have perceived) what is not (or was not) really perceived. We may mistake ‘falsehood for demonstration’ by failing to examine presuppositions. But a false proposition can only be entertained as a ‘belief’ and in virtue of an act of ‘judgement’: ‘Knowledge being to be had only of visible certain Truth, *Error* is not a Fault of our [faculty of] Knowledge, but a Mistake of our [faculty of] judgement giving Assent to that, which is not true.’⁵⁶ We can, it seems, be guilty of mistaking presumption for perception, and what is extrinsically determined for what has intrinsic evidence. Yet Locke’s discussion of this sort of mistake is unsystematic and largely incorporated into a general explanation of error which concentrates on the different topic of our frequent failure to judge in accordance with probabilities. Its interest lies chiefly in the light it casts on his conception of ‘evidence’ or what might be called the intrinsic perspicuity of knowledge. That too, will be considered and explained below.

Other divisions of knowledge

Locke employed the word 'knowledge', like the expression 'to have an idea', with linked occurrent and dispositional senses. It is in the occurrent sense that knowledge is defined as the perception of a relation between ideas: '*actual Knowledge...is the present view the Mind has*' of such a relation.⁵⁷ An objection that the ordinary English word 'knowledge' and its synonyms in other languages have only a dispositional sense is itself without much force. 'Perceive' and 'see' (as in 'I saw that *q* followed from *p*') are primarily occurrent verbs, and 'understand' can either be occurrent or dispositional. To deny that 'know' can intelligibly be given an occurrent sense like that of 'understand' which is appropriately linked to its normal dispositional sense is in effect to do no more than register disapproval of anything that smacks of intuitionism. It is, indeed, significant that disapproval of intuitionism has often found expression, in recent discussions, in the paradoxical claim that even in such a sentence as 'I understood the proof when I read it' the verb 'understand' is really dispositional, attributing a capacity.⁵⁸ But just because that claim is paradoxical, it supplies little reason to reject Locke's conception of 'actual knowledge' out of hand.

Because Locke was an intuitionist, for whom 'evidence' was a condition of knowledge, he placed limitations on the extent of dispositional knowledge which not all philosophers would accept. It is not enough for someone to have known a truth dispositionally that he accepts it as soon as it comes to be considered.⁵⁹ The acceptance must be due to its 'having been once laid before his Thoughts', so that 'he evidently perceived the Agreement, or Disagreement of the *Ideas* whereof it consists.' A man's habitual knowledge must be 'lodg'd in his Memory, by a foregoing clear and full perception.'⁶⁰ Perception is necessarily conscious, and without perception there is no understanding or grasp of the truth, and so no knowledge or 'science'. That point seems to lie at the heart of Locke's rejection of the theory of innate dispositional ideas and knowledge. It is certainly not true, as Leibniz and many later critics have implied, that he never properly considered the theory.⁶¹ His view was that there is no knowledge without evidence, and no evidence without actual consciousness.

Locke had a rather sophisticated conception of memory, although it apparently did not occur to him until after the publication of the third edition. It is of interest

because of its relations both to a problem in Cartesian philosophy and to a twentieth-century debate. 'Habitual knowledge' is itself divided: in one sort of case, when the remembered truth is revived any necessary proof is revived too, and we actually perceive the truth again. In the other case we remember the conclusion with conviction, but not the proof. In this second kind of remembrance, so Locke proposed in the earlier editions, the subject strictly speaking 'rather believes his Memory, than knows the thing; or rather it is something between Opinion and Knowledge, a sort of Assurance which exceeds bare Belief, for that relies on the Testimony of another'.⁶² One difficulty with this odd account is the problem of how there could be something *between* belief and knowledge, given Locke's explanation of the distinction. Given Locke's model, it would seem that an 'affirmation' is either based on extrinsic and contingent considerations or it is not. If not, then it must either be immediate or be based, like 'demonstration', on considerations intrinsic or necessary to the terms in relation. Memory, however, seems in itself immediate in its deliverances and so to that extent classifiable with intuition and sense, where Hobbes placed it. If it is said that, even so, memory delivers less than knowledge because the subject has to *trust* its *testimony*, that introduces an element into the account of cognitive faculties contrary to the whole spirit of Locke's epistemology. Intuition and sense can just as well (or as mistakenly) be regarded as giving 'testimony' which we need reason to 'trust'. The argument would thus reduce all knowledge to belief.

A second difficulty relates to degrees of 'evidence'. If, as Locke and Descartes both suggested, demonstration is less 'evident' or certain than intuition partly because memory is involved in long proofs, then the first edition account of memory threatens to reduce demonstrative knowledge to something less than knowledge. The distinction drawn between types of habitual knowledge (depending on the presence or absence of an ability to revive the proof in order as a whole) is irrelevant, since memory is said to have a role *within* the particular occurrence of the proof. As Locke pointed out in the fourth edition, there must be something wrong with a view which requires that Newton either holds the whole of *Principia* in mind at once or does not know that its conclusions are true.⁶³ Even if memory were admitted as a separate faculty delivering immediate knowledge, a similar problem would remain. For memory knowledge would then have to be a separate degree of knowledge, presumably lower on the scale than sense. Thus the second degree of knowledge would be reducible to a fourth.

Locke consciously recognized at least the second difficulty, and his radical and reasonable solution was in effect to deny that memory is a distinct cognitive faculty at all: rather 'remembrance is but the reviving of some past knowledge'. The possessor of habitual knowledge retains in memory the evidence or certainty of a proposition even if the proposition is not now evident. He cannot doubt that the truth which he once perceived still holds.⁶⁴ Sensitive knowledge too can be retained: a witness 'knows such a man wounded another remembering that he saw him run him through'. It is amusing (and instructive) that twentieth-century

philosophy has re-enacted something like Locke's change of mind about memory. In one discussion Russell treats memory as an intuitive faculty giving direct knowledge of the past; in another, more sceptical mood, he presents it as a source of images which supply the content of somewhat insecure beliefs about the past. But more recently Ryle and others have brought us round to a 'retained knowledge' view, with its many theoretical advantages.⁶⁵

Another main division of knowledge is with reference to the types of proposition known. Locke tells us that the kinds of 'agreement or disagreement' between ideas can be reduced to four. They are given as:

- (1) Identity or diversity
- (2) Relation
- (3) Coexistence or necessary connection
- (4) Real existence.⁶⁶

Many of the issues raised by this classification must be left until later chapters, but some explanation can be given here.

'Identity or diversity' is the only one of these classes of propositions all members of which are perceived intuitively. That is because on Locke's view the perception of the identity or diversity of ideas reduces to the 'discerning' or distinguishing of a single idea or two different ideas. Since we cannot have ideas without distinguishing them, the potentiality for a kind of propositional knowledge is bound up with the very having of ideas.⁶⁷ Yet Locke stressed that, as explicit propositional knowledge, knowledge of identity is entirely 'trifling' and useless 'for the enlightening the Understanding in any part of knowledge'.⁶⁸ Even that much of his argument makes it clear that he was concerned, not with identity propositions in general, which can often be highly informative, but with those which are conceptual or in a strong sense tautological, e.g. 'Round is not square', 'A vacuum is a vacuum', or 'that most general one, *What is, is*'.⁶⁹ Such propositions as 'The accused is (identical with) the culprit' are very differently treated in the chapter 'Of Identity and Diversity'.

The second class is 'relation' between distinct ideas, for example, 'Two Triangles upon equal Basis, between two Parallels are equal.' This class of propositions may seem surprising, since it seems to be a ragbag capable of including others. Locke himself confesses that 'Identity and Co-existence are truly nothing but Relations', although 'they are so peculiar...that they deserve well to be considered as distinct Heads'.⁷⁰ Yet closer inspection reveals that 'relation' is not a mere 'miscellaneous' class, but is deliberately opposed in its own right to the first class as well as to the third. Indeed this opposition would explain why 'identity or diversity' is separated out at all, for Locke seems to have been motivated by something more than an eye for superficial peculiarity.

One of his intentions was clearly to denigrate 'maxims', and in particular the law of identity. Yet he did not merely denigrate this logical law, for his position is in effect as follows: reasoning which is logically *in accordance with* the law of

identity, and which would therefore not be significantly altered were the law explicitly introduced as a premise or principle, does not (at any rate prior to such expansion) in any sense employ or even rely implicitly on the law of identity, but rather involves the simple exercise of a primitive and unanalysable faculty of 'discernment'. To possess that faculty and the other such mental faculties which constitute human reason is not to possess innate knowledge of a set of very general logical principles, even though it is to possess what is required to see the truth of those principles, among other less general truths, once they are understood. Similarly it seems to have been Locke's official view that to discriminate or identify ideas (i.e., for example, to perceive the truths expressed as 'White is not black' or 'White is white') is not actually to employ the ideas of diversity and identity. For these latter, highly abstract ideas are not any part of the equipment of children who can perceive such particular truths intuitively and without difficulty, simply in virtue of their capacity of discernment.

This argument is one of a number of analogous lines of thought in the *Essay* which together form an interesting and significant element in the campaign against innate ideas and knowledge. To take another example already considered above, Locke held that simply to think involves using ideas as natural signs. The fundamental use of an idea as a sign is the use of a simple idea to represent or stand for its regular cause, a power or 'quality' in things. It might therefore appear that we cannot use ideas in thought unless and until we possess the ideas of *cause* and *power*, and at least implicitly accept the principle, 'All our sensations have causes'. Yet Locke seems rather to have assumed that the ability to think about things is primitive, and therefore not something consequent upon an antecedent conception of causality or an innate principle that every event has a cause. For him that general principle can be formulated and known only after the general ideas of power and cause have been explicitly formed in response to experienced regularities. To say that it was already implicitly present or known would be to confuse the capacity to use the natural language of thought with the possession of explicitly formulated propositional knowledge. In general Locke held, in opposition to a considerable tradition, that the capacity to think reasonably is in no way attributable to an antecedent set of implicitly known principles of rationality. Something like this controversy has recently emerged in the question how far the acquisition of language is to be ascribed to innate knowledge of its fundamental rules or 'deep structure', rather than to general intelligence and pre-linguistic cognitive capacities.⁷²

Yet the interest and peculiarity of Locke's first type of proposition, in contrast with the second, probably extended in his mind further than his estimation of maxims. In the chapter 'Of Trifling Propositions', he in effect enlarged the first class to include, not only 'purely identical propositions' but all propositions true by definition, even those which predicate only a 'part of the Definition of the Term defined', e.g 'Lead is a metal' or, on a certain definition of 'gold', 'All gold is fusible.' The notion of an 'identical proposition' was itself problematical, as Locke noted, for some people would include definitions.⁷³ And, as with

definitions, the introduction of part-definitions into an argument, so he claimed, 'can only serve to show the Disingenuity of one, who will go from the Definition of his own Terms, by re-minding him sometimes of it; but carry no Knowledge with them, but of the Signification of Words, however certain they be'.⁷⁴ Once again Locke may have been hitting out in more than one direction, in this case against more than one conception of analytical method. First, and most explicitly, he was arguing against the Aristotelians that, if in examining the attributes of a species such as *man* we come across a number of common properties which seem hard-core, necessary or essential, that is not because we are uncovering and coming to understand the true nature of the thing itself. On the contrary, we are merely explicating what we mean by 'man', the simple ideas which go to make up our complex idea and determine what for us is to count as a human being. The argument was a plank in an anti-Aristotelian platform of (as we shall find) considerable complexity.

Another important target is likely to have been Hobbes, for whom all science, whether natural, mathematical or political, starts from definitions: 'all necessary propositions are either definitions, or parts of definitions, or follow from definitions'.⁷⁵ Judicious definitions, arrived at after a proper analysis of experience, serve, Hobbes thought, as the explanatory principles from which, together with particular hypotheses, particular *explicanda* are synthetically deduced. For Locke, however, definitions may clarify our meaning but otherwise play no significant role in advancing knowledge. It is false that all necessary truths are analytic, for science requires another sort of necessary proposition:

we can know the Truth, and so may be *certain* in Propositions, which affirm something of another, which is a necessary consequence of its precise complex *Idea*, but not contained in it. As that the *external Angle of all Triangles, is bigger than either of the opposite internal Angles*; which relation of the outward Angle, to either of the opposite internal Angles, making no part of the complex *Idea*, signified by the name Triangle, this is a real Truth, and conveys with it instructive *real Knowledge*.⁷⁶

So with the joint purpose of explaining the informativeness of mathematics and the barrenness of a purely analytic method in natural philosophy, Locke drew a distinction which was the source of Kant's better-known dichotomy between analytic and synthetic propositions. Now the given instance of an instructive necessary (i.e. synthetic *a priori*) proposition is the same in kind as the instance of 'relation', while the chief point of making a separate class of affirmations of 'identity or diversity' is to stress their trivial or verbal character. It is therefore difficult not to conclude that Locke understood the distinction between his first two classes of proposition as in effect a distinction between analytic and synthetic necessary truth.⁷⁷

The issues raised by 'coexistence or necessary connection' are also closely related to the trifling-instructive distinction, although this category of

propositions has solely to do with substances and natural philosophy. On Locke's account the proposition 'White exists' or 'Something white exists' falls under 'existence', and can only be known to be true by sense-perception. 'A horse exists', however, falls under 'coexistence', since it is taken to be equivalent to the proposition that the set of observable attributes by which we define the name 'horse' coexist in, or are united in, the same substance or thing. Observable attributes are united in a material thing or substance, whether particular or general, when the thing is so constituted or structured at the most fundamental physical level as to have those attributes together. For example, the constitution of gold, or of a piece of gold, causes it to have great weight, yellowness, malleability and solubility in *aqua regia*, as well as many other attributes. Locke held, as we have seen, that if we define 'gold' by this particular list, then the universal predication 'Gold is malleable' is necessary but trifling, falling into a class at least closely allied to 'identity'. If, on the other hand, some non-defining attribute is predicated of gold, say solubility in selenic acid, then the proposition is informative and falls under 'coexistence', since it affirms that this attribute coexists with the defining attributes.

How can a true 'coexistential' proposition be known to be true? Locke's position seems to have been as follows. If the proposition is particular, it can be known through the senses; as we shall find, this assumption raises certain difficulties. Where the proposition is universal, however, its truth cannot be known simply through experience, since the coexistence of attributes in all observed instances does not entail universal coexistence, even if it makes it probable. The proposition could only be known to be true if we could perceive that the attributes necessarily coexist.⁷⁸ Such perception is impossible for us, since all we know of substances are their observable attributes, i.e. the powers which they have to affect observers through the senses, and observably to affect other substances. Locke believed, however, that if we knew the fundamental constitutions of substances we could in principle demonstrate that what has such and such observable attributes *must* have this further attribute. The necessity could then be 'perceived', Locke thought, in the way in which we can 'perceive' that a figure of such and such a shape must have such and such other geometrical properties. Thus knowledge of the universal 'coexistence' of attributes and knowledge of their 'necessary connection' come to the same thing. Yet 'coexistence' and 'necessary connection' are clearly not in general the same, since attributes may coexist contingently. Some heavy and malleable metal is yellow, but some is not.

Locke's conception of the possibility, at least in principle, of our perceiving necessary connections between natural attributes is part and parcel of his mechanistic view of the material world. Roughly, he held that the unknown ultimate laws of motion flow necessarily from the very being or essence of matter as space-occupant. It is our ignorance of this essence and of these principles, together with our ignorance of the ways in which matter is structured in particular material substances, which denies us physical 'science' and limits us, as students

of nature, to 'natural history' or the recording of coexistences within our experience. The more carefully we base our classification of substances on systematic experiment and observation, however, the more probable it is that we shall formulate coexistential propositions, such as 'Gold is soluble in selenic acid', which are true, expressing what is in fact a necessary connection, even if not to us a 'visibly' necessary connection. The interpretation of Locke on this issue calls for justification, and some will be given in Volume II. It will also be necessary to consider the shortcomings of his version of mechanism, and the question of how far his model for the relationship between science, 'natural history' and classification could survive its demise.

Propositions falling under 'existence' are known in three ways. That a thinking thing exists in oneself can be known by intuition, and 'our very Doubts about what it is, confirm the certainty of its being'.⁷⁹ That an eternal, necessarily existent and necessarily thinking thing exists can be known by demonstration. Otherwise we have only sensitive knowledge of existence. 'Existence' is the only category of proposition which Locke did not mention as a possible sub-class of 'relation'. That is presumably because he saw the other categories as concerned with the inter-relations within a network of ideas, while existential propositions are concerned with the relationship between the network and the world. Yet he nevertheless explained the perception of an existential truth as the perception of the agreement of some idea with the idea of existence, in accordance with his general definition of knowledge. To all these points we shall return.

Probability and the nature of ‘assent’

Locke’s approach to ‘belief’ and to probability, like his theory of ideas, is complicated by his employment of models which he evidently took to be compatible but which are not easily reconciled with one another. The most fundamental model, from the point of view of his general theory, derived from his conception of ‘affirmation’ as a mental act of the combination of subject-concept and predicate-concept. It is embodied in the first official definition of ‘judgement’, as opposed to ‘knowledge’, to appear in the *Essay*.

Thus the Mind has two Faculties, conversant about Truth and Falshood. *First, Knowledge*, whereby it certainly perceives, and is undoubtedly satisfied of the Agreement or Disagreement of any *Ideas*. *Secondly, Judgement*, which is the putting *Ideas* together or separating them from one another in the Mind, when their certain Agreement or Disagreement is not perceived, but *presumed* to be so...And if [the mind] so unites, or separates them, as in Reality Things are, it is *right judgement*.⁸⁰

‘Probability’ is ‘the appearance of such an Agreement or Disagreement’ by the intervention of persuasive, but not deductively certain ‘proofs’.⁸¹ The combinatory act of ‘assent’ has degrees related to the various ‘grounds of probability’, which grounds are both ‘the foundations on which our Assent is built’ and ‘the measure whereby its several degrees are, or ought to be *regulated*’.⁸² The degrees of assent are ‘such different Entertainment, as we call *Belief, Conjecture, Guess, Doubt, Wavering, Distrust, Disbelief, etc.*’.⁸³ ‘Wavering’ recalls Hobbes’ account of doubt as alternating opinion, and it seems clear that, at least on this model, all the items on Locke’s list are supposed to be propositional attitudes of the same general kind.

A fundamental problem with the model is that it lacks the means of distinguishing such and such a degree of assent to the proposition p from unqualified assent to the proposition that on the available evidence p is probable to such and such degree. In other words, it excludes the possibility of what seems to be an essential precondition of our possessing a notion of objective probability at all, i.e. our keeping a proposition neutrally in mind as a hypothesis to be set

against the grounds for believing it. In 1671 Locke had apparently favoured a traditional approach to this difficulty for the view of propositions as affirmations, at the same time echoing Hobbes on 'faith'. Distinguishing between knowledge and faith, he had concluded that with respect to the latter 'there is always a necessity of words or signs', since we 'first hear the words and afterwards examine the thing or truth of them'.⁸⁴ It is not very surprising that this line of thought was suppressed, not only because there is more to probability than the assessment of testimony, but because a concession that we can 'examine the truth' of a sentence otherwise than by first considering the truth of the propositional thought which it expresses would be violently at odds with Locke's mature conceptions of truth and meaning.

Nevertheless, in favour of his earlier view, it does seem that the conditions under which an animal or infant can truly be said in some sense to 'entertain the proposition that *p*' will always include something like an awareness that *p*, or belief that *p*, or fear and suspicion that *p*. Beings without language have beliefs, but they do not have that capacity to comprehend what a belief is, or to identify and distinguish particular possible beliefs, which is implicit even in the simple capacity to ask a question. A being with the capacity to express its beliefs in language, however, must *ipso facto* have the notion of an assertion such as anyone might make, identifiable by means of the same sentence as could be employed by anyone to make it (allowing for the context-dependent role of demonstratives, personal pronouns, tense and so forth). The acquisition of that notion of same-saying is an essential part of the learning of language and is self-evidently impossible without it; and it is arguably only through that notion that anyone could arrive at the notion of a particular proposition or belief. In that case the neutral apprehension of a proposition is essentially an apprehension of meaning, and so dependent on the possession of language.⁸⁵

However that may be, it seems that on Locke's official view a judgement of probability is not a second-order judgement. Yet he also seems often enough to have assumed that that is just what it is. Certainly he identified some sort of objective probability as 'the proper Object and Motive of our Assent'.⁸⁶ In some passages he even seems in danger of collapsing 'opinion' into 'knowledge' by allowing, despite all, that we 'perceive' probabilities, although it is worth noting that even here it is the connection between subject and predicate which we are said to perceive, not (as any modern notion of 'perceiving probabilities' would surely presume) the connection between a hypothesis and the evidence for it.⁸⁷ Another, perhaps rather clearer manifestation of a view of probability-judgements as 'second-order' occurs with the suggestion that, at least in some cases, we can 'suspend' judgement while searching out and examining the available grounds for some opinion.⁸⁸ But how far that suggestion is in conflict with the general theory that some sort or degree of 'affirmation' supplies the propositional bond depends, of course, on how far Locke would have regarded suspense of judgement as itself a degree of assent, perhaps identical with 'waving' or 'doubt'. To this question I will shortly return.

These and similar tensions reveal themselves in particular in Locke's account of error. A major problem for any approach like his is that it makes 'judgement' a cognitive faculty, a way of being informed about the world, and yet at the same time makes every belief, however wild and unreasonable, by its very nature as a belief an exercise of judgement. The problem lies partly in the close comparison with knowledge. Judgement is one of the 'understanding Faculties', which has been 'given to man to supply the want of clear and certain knowledge in Cases where that cannot be had'.⁸⁹ As knowledge is correspondent to 'evidence', so belief is correspondent to probability. Belief, that is to say, is naturally (almost, it seems, by definition) founded on the sort of extrinsic 'arguments or proofs' which are the grounds of probability.⁹⁰ As John Passmore says, in perhaps the only sustained discussion of Locke's theory in this respect, it seems that 'Locke defines belief...as a purely intellectual operation, in a way which emphasizes at once its likeness, and its inferiority to knowledge.'⁹¹ Yet it is the same faculty of judgement which is responsible for error and mistake. Its weakness does not lie simply in the weakness of probability as opposed to certainty, or in the fact that what is probable may be false. Unlike the knowledge-delivering faculties of intuition and sense, or the knowledge-preserving faculty of memory, judgement can be perverse, wrong-headed and irrational. If it is a cognitive faculty, then, as Locke had to recognize, it is one which can be abused, or at least can be misguided by non-rational influences. Passmore argues that there is a fundamental incoherence in Locke's thought about belief centred on his ambivalent attitude towards the question whether belief is voluntary and so in itself a proper object of praise or blame. Eventually, Passmore believes, Locke in effect abandoned his original 'intellectualist' conception of belief as the exercise of a cognitive faculty, swinging round to a view of it as, like action itself, always motivated by passions of one kind or another, whether disreputable or reputable.

There is no doubt that the tone of much of Locke's account of belief is moralistic, sometimes intensely so, but that is also true of some of the things he had to say about knowledge, or rather about our failure to acquire it when it is to be had. The question is whether his morality of rational belief and error takes essentially the same form as his morality of knowledge and ignorance. As far as knowledge is concerned, it is clear that he totally rejected the Cartesian doctrine, discussed in [chapter 2](#) above, that a voluntary act of assent follows, and is distinct from, a clear and distinct perception or idea. Knowledge for Locke simply *is* perception, and the relevant analogy is with sight. We may choose whether we turn our eyes towards certain objects and whether, having done so, we 'curiously survey' them, but we cannot choose what we then see. Similarly

all that is *voluntary* in our Knowledge, is the *employing*, or withholding any of *our Faculties* from this or that sort of Objects, and a more, or less accurate survey of them: But they being employed, *our Will hath no Power to determine the Knowledge of the Mind* one way or other.⁹²

It certainly seems that Locke wanted to extend this un-Cartesian separation of the provinces of will and understanding (or 'mind') to the case of belief, for he finally concluded, all things considered and due qualifications made, that, 'As Knowledge, is no more arbitrary than Perception: so, I think, Assent is no more in our Power than Knowledge.... And what upon full Examination I find the most probable, I cannot deny my Assent to.'⁹³ On this model, the place at which moral considerations and praise or blame can be brought to bear is not between 'full examination' and assent, but at the earlier stage: '*we can hinder both Knowledge and Assent, by stopping our Enquiry, and not imploying our Faculties in the search of any Truth. If it were not so, Ignorance, Error, or Infidelity could not in any Case be a Fault.*'⁹⁴ It seems then, that Locke postulated a voluntary power which we possess prior to (but not after) 'full examination': the power either to suspend belief for the purpose of more enquiry, or to stop enquiry in order to stick with the current 'apparent probability'. Blameworthy error, it is suggested, results from the wrong employment of this power.

The situation in which the power exists is elsewhere described more positively:

in Propositions, where though the Proofs in view are of most Moment [i.e. *scil.* weigh more heavily in their favour], yet there are sufficient grounds, to suspect that there is either Fallacy in Words, or certain Proofs, as considerable, to be produced on the contrary side, there Assent, Suspense, or Dissent, are often voluntary Actions.⁹⁵

Yet there are various problems with this whole proposal. For one thing, the analogy with sight and knowledge does not bear much probing. If 'full examination' corresponds to looking (or looking closely) and any degree of assent corresponds to seeing, then suspending assent must correspond, absurdly, to choosing not to see in order to look (or to look more closely).⁹⁶ And since choosing to 'stop enquiry' and settling for some degree of assent may come to the same thing (and were so treated by Locke), refraining from looking ought on this model to be equivalent to seeing. It could perhaps be argued that choosing not to look closely at something glimpsed is equivalent to choosing to see it no better than superficially, but someone who 'stops enquiry' (unless from lack of interest or the like) may be choosing to *believe* something, not to leave it in limbo. A difficulty of another kind is that grounds for suspecting that an argument for *p* is fallacious or incomplete seem themselves to be reasons or 'proofs in view' which should affect the 'apparent probability' of *p*. If they have their due effect on us, and we have the appropriate doubts, then we may be inclined (if the matter is interesting and important enough, and there is time) to investigate and reflect further. But it is not obvious how the existence of such grounds is supposed to make us free to *believe* or *disbelieve* the original proposition. And if we *are* peculiarly free so to ignore reasons of this type, that in itself might seem to be a counter-example to the general model according to

which the degree of assent (*unless* we suspend assent) is determined by the apparent degree of probability.

Nevertheless Locke's position is not desperately mysterious or odd. First, there is, after all, a continuum of judicious thought, at one end of which the metaphor of perception and sight is most attractive, while at the other we talk naturally enough of decision and choice. It is at least to some extent plausible that the language of decision and choice is most appropriate when the credentials of stated or given evidence are themselves in question. Choosing to believe or disbelieve a witness is perhaps a paradigm case. Second, Locke described doubts about the validity of an argument and doubts about its completeness or balance, given that they have any justification at all, as 'two ways...of evading the most apparent Probabilities'. He evidently felt that this sort of second-order argument is peculiarly prone to manipulation, so that slight grounds for doubt can be given undue weight: it 'is a refuge against *Conviction* so open and so wide, that it is hard to determine, when a Man is quite out of the Verge of it'.⁹⁷ At the same time he argued in defence of tolerance that it is unreasonable to expect people to be satisfied even by clear and decisive arguments before they have had time to reflect on them.⁹⁸ His position can seem to reduce to the claim that, within limits, it is up to us how *cautions* we are. 'I know not yet all that may be said on the contrary side'⁹⁹ can after all be more of an expression of caution than an argument. Blame attaches both to those who are too cautious (like tobacco manufacturers who, knowing the evidence, nevertheless suspend belief on the question as to whether smoking causes cancer) and to those who are not cautious enough (as well as to those who, as they say, 'don't want to know'), i.e. both to those who refuse to recognize when 'fair examination' has been made, and to those who refuse to make it.

What is basically wrong with Locke's model is its firm separation of two stages, the first voluntary, the second determined by ground or reason for belief. In a way he seems to have recognized this weakness, in that he did not himself adhere strictly to the model. The decision to 'stop enquiry' tends to coalesce, as we have seen, with a decision to believe or disbelieve, so that 'Assent, Suspense and Dissent' are put explicitly on the same level.¹⁰⁰ This coincides with the ambivalent status of 'suspense', which needs sometimes to be presented as an act antecedent to any judgement, and yet at others seems itself to be a judgement or 'degree of assent' (the second interpretation being the more consonant with the general theory of 'affirmation' as the putting together of ideas). But the inadequacy of the two-stage model appears as clearly as anywhere in the context in which it is most firmly stated, with the itemization of the sources of error and disagreement.

Locke distinguished four types of case: in order, error can be ascribed to 'want of proofs', 'want of ability to use them', 'want of will to use them' and 'wrong measures of probability'.¹⁰¹ 'Want of proofs' comprises unavoidable lack of evidence which is practically available, and is ascribed to the necessities of living, in the case of the poor, or to the narrowness and intolerance of society.

The topic was dear to Locke's heart, reflecting both his consuming concern for religious toleration and freedom of thought, and his need to feel that everyone has (or, if not, should be given) enough leisure 'to think of his Soul, and inform himself in Matters of Religion' and morality.¹⁰² The second source of error, incapacity to 'discern that side on which the strongest Proofs lie', is significantly different in that it has its effect *between* enquiry and judgement. Locke expressly declined to speculate as to the physical basis or causality, in nature or nurture, of the 'great difference in Men's Intellectuals'; but it seems that the faculty of estimating probabilities is itself capable of malfunction while still describable as 'reasoning'.¹⁰³ The third source of error, the lack of will to find and use evidence, includes mere laziness but is also ascribed to such motives as fear and self-interest. It is like the first in constricting inquiry, but is different from both the first and the second in being characteristically blameworthy.¹⁰⁴

The fourth source of error, however, comprising 'wrong measures of probability', both is often blameworthy and comes, or at any rate seems designed to come, between inquiry and judgement. Its victims believe in the face of the evidence, and sometimes even in the face of the deliverances of the senses,¹⁰⁵ but not through intellectual incapacity. Manifest probabilities may be outweighed by prejudices (ingrained 'by long Custom and Education', sometimes with the help of pride), by 'Men's Appetites, and prevailing Passions', and by 'common received Opinions' (especially those of 'a Party that Education or Interest has engaged them in').¹⁰⁶ The striking feature of this list is its implication that our estimate of probability itself, and not just the decision whether to suspend assent or to stop inquiry, can be directly determined by the interests, appetites and passions which motivate action. Since our estimate of probability is hardly to be disentangled from a degree of assent,¹⁰⁷ it would follow that assent may in itself be blameworthy. Since that would be the end of the two-stage model and the analogy with knowledge, it is perhaps not surprising that Locke struggled to disown the consequence:

Not but that it is the Nature of the Understanding constantly to close with the more probable side, but yet a Man hath a Power to suspend and restrain its Enquiries, and not permit a full and satisfactory Examination, as far as the matter in Question is capable, and will bear it to be made. Until that be done, there will be always these *two ways left of evading the most apparent Probabilities*.¹⁰⁸

These 'two ways', the appeal to doubts about an argument's validity and the appeal to doubts about its completeness, have perhaps been sufficiently discussed. But the logic of this passage would suggest an oddly complicated and barely intelligible story: that a passion or interest might motivate us to restrain inquiry, thus giving us the chance to appeal to the (perhaps remote) possibility of a presented argument's being fallacious or incomplete in order to arrive at a judgement (which might itself be suspension of belief) contrary to an otherwise

‘apparent probability’. Which stages of such a convoluted process might be voluntary, and which not, is obscure. Locke himself evidently felt that he had somehow vindicated the simple two-stage model, which he then proceeded to expound at some length. Yet if he *had* succeeded in doing that, it would be difficult to see why interest and the passions have a place in the discussion of ‘wrong measures of probability’ at all.¹⁰⁹

Passmore draws on such difficulties as these in order to present a picture of Locke’s being gradually weaned away from an ‘intellectualist’ model, according to which all belief is involuntarily motivated by reasons and probability, towards a view of it as characteristically directly motivated by passions and inclinations and therefore, by and large, voluntary. The chief evidence for this alleged ‘new emphasis on passion’ he finds in the chapter ‘Of Enthusiasm’ added to the fourth edition, which he takes to illustrate how ‘the existence of the enthusiast constantly undermined Locke’s hopeful view of man, as a being who naturally, and inevitably, responds to the most probable hypothesis’.¹¹⁰ Here Locke described even ‘rational belief not in terms of a purely intellectual weighing-up but rather in terms of the operations of a certain form of passion—the love of truth’. Passmore draws the conclusion that Locke had given up his initial presumption of the natural rationality of man from Locke’s statement that for only the ‘very few lovers of Truth for Truth’s sake’ does their degree of assent match the degree of probability. In all other cases the mismatch is due to ‘some other Affection, and not to the love of Truth’. Belief which goes beyond its justification, ‘is owing to our inclinations that way, and is so far a Derogation from the Love of Truth as such’.¹¹¹ All this, Passmore claims, ‘entirely disrupts’ the argument previously developed in the surrounding chapters. Locke would have liked to believe that human beings are automatically rational, going wrong only when some of the evidence is not before them, but he could not ultimately reconcile this thesis with his actual experience of irrationality. Influenced, ‘more particularly, by the “enthusiasm” of the Puritan sectaries’ he was ‘gradually driven into a new picture of belief, in which it is no longer a weaker form of knowledge, but rather...an attempt to remove uneasiness, to satisfy our inclinations’.

It is difficult to see that this rather startling interpretation can be justified. By dealing with the claims of immediate revelation to be an independent ground of assent, ‘Of Enthusiasm’ filled a rather obvious gap left in the previous chapter, where discussion is almost exclusively confined to ‘traditional’ revelation. It anticipates the more general discussion of error in so far as there is need to explain away the apparent evidence and subjective certainty of the enthusiast’s beliefs. The rhetorical contrast in the opening paragraph between the love of truth and other ‘affections’ is perhaps a rather loose way of introducing the topic of the explanation of error, but it hardly derogates from rationalism or ‘intellectualism’ in Passmore’s sense. Rather like Kant’s ‘sense of duty’, the ‘love of truth’ is a respect for reason for its own sake, not itself one of our ‘inclinations’ or ‘passions or interests’, but explicitly opposed to them. The

enthusiast lacks it, whether or not his beliefs are true. Those beliefs themselves Locke went on to ascribe, not to the conventional passions or desires (the enthusiast's beliefs are not classed with the lover's belief that his mistress is true),¹¹² but to the imagination. They are 'the ungrounded Fancies of a Man's own Brain'. Enthusiasm rises 'from the Conceits of a warmed or over-weening Brain', it sets up 'phancy for our supreme and sole Guide'.¹¹³ As such, it is an 'internal impulse' which

like a new Principle carries all easily with it, when got above common Sense, and freed from all restraint of Reason, and check of Reflection, it is heightened into a Divine Authority, in concurrence with our own Temper and Inclination.¹¹⁴

'Inclination' and the conventional 'passions' enter this story at a rather late stage: because it 'so flatters many Men's Laziness, Ignorance and Vanity...when once they are got into this way of...certainty without Proof, and without Examination, 'tis a hard matter to get them out of it'.¹¹⁵ It is not an unreasonable passion which was Locke's fundamental target, but a misconception; and his argument runs on familiar enough lines. When men believe themselves divinely inspired, 'Does it not stand them upon, to examine upon what Grounds they presume it to be a Revelation from GOD?'¹¹⁶ The passions get in the way of this duty of examination, but are not responsible for the original conceits. If there was anything new in the fourth edition explanations of error in both 'Of Enthusiasm' and 'Of the Association of Ideas', it lay in the more conspicuous role of the physiology of the imagination. Even that had been anticipated in a journal entry of 1682.¹¹⁷

We need not, then, suppose that Locke experienced any sweeping conversion to irrationalism. There are inconsistencies and vacillations in his account of judgement and error, but it is arguable that they stem from difficulties inherent in his subject-matter which were not (and are not) to be evaded by any such precipitate change of direction. So, at least, it will be argued below.

The grounds of probability

Locke's account of the grounds of assent or probability should be read with recognition of his concern with three things: the programme of 'natural history', the status of speculative hypotheses and the status of religious authority and faith. He began with two main distinctions. The first is between, on the one hand, 'the conformity of any thing with our Knowledge, Observation and Experience' and, on the other hand, the 'Testimony of others'.¹¹⁸ The second distinguishes propositions concerning particular 'matters of fact' falling within human experience and so capable of human testimony, from propositions concerning what lies 'beyond the discovery of the senses', and so beyond the scope of human testimony.¹¹⁹

The probability of 'particular matter of fact' depends on the 'certainty of observations', and the 'frequency and constancy of Experience', both of ourselves and others.¹²⁰ Although Locke concentrated on the relevance of general experience and report to judgements about particulars, what he said evidently applies also to generalizations. Thus the highest 'degree' of probability is when the general consent of others concurs with the subject's constant experience: 'such are all the stated Constitutions and Properties of bodies, and the regular proceedings of Causes and Effects in the ordinary course of Nature'.¹²¹ The second 'degree' is when experience and testimony suggest that a thing is for the most part so.¹²² The third, when unsuspected witnesses report, without contradiction, something which experience allows might as well have been so as not.¹²³ Finally, it can happen that 'the reports of History and Witnesses clash with the ordinary course of Nature, or with one another'. The probability 'rises and falls, according as those two foundations of Credibility... favour or contradict it'. Here, he says, there are no 'precise rules'.¹²⁴

In fact none of Locke's discussion suggests any precise calculus, or even the now so familiar thought that the first two 'degrees' of probability might themselves be thought capable of degrees: the first, in accordance with (let us say) the number, or number and variety, of like cases observed; and the second, according to the relative frequency of a particular outcome among like cases. It seems that Locke wished to do little more than identify 'natural history', whether formally or informally, individually or corporately done, as the fundamental

source of reliable or fairly reliable, and therefore useful information. Here we can draw a contrast with anything like Russell's 'Principle of Induction', which, although a direct descendant of Locke's discussion via Hume, is intended as a principle of rationality or a quasi-logical measure of probability that is independent of ontological presuppositions.¹²⁵ For Locke, on the other hand, absolute or relative regularity in our experience is a ground of *belief* just because it is some indication, always in principle fallible, of necessities which might otherwise be *known* (if not by imperfect human beings): 'For what our own and other Men's constant Observation has found always to be after the same manner, that we with reason conclude to be the Effects of steady and regular Causes, though they come not within the reach of our Knowledge.'¹²⁶ Another controversial issue hinted at in Locke's treatment of probability as to 'matters of fact' is that of miracles, to which I will shortly return.

The category of unobservables, which is opposed to 'matters of fact', includes spirits and very remote or very small material things. Most importantly for Locke's philosophy of science, it includes the underlying causes of the regular effects we observe, and their 'manner of operation'. When we speculate about unobservables, analogy with what falls within our experience 'is the only help we have' and the only ground of probability. For example, finding that heat and fire can be produced by friction, and that a variation in the surface-texture of such things as silk and velvet will alter their colours, we reasonably speculate that heat and fire consist quite generally in the violent motion of particles, and that different colours are always to be explained by the different refraction of light from different surface-textures:

This sort of Probability, which is the best conduct of rational Experiments, and the rise of Hypothesis, has also its Use and Influence; and a wary Reasoning from Analogy leads us often into the discovery of Truths, and useful Productions, which would otherwise lie concealed.¹²⁷

Such remarks should be set beside Locke's other references to the hypothetico-deductive method, of which Descartes' and Hobbes' conceptions of analysis and synthesis, discussed in [chapter 11](#) above, are forms. One significant historical question concerns Locke's attitude towards the 'inductive' method of Bacon, as set out in what were among the most celebrated seventeenth-century works on method. Bacon's argument stood behind the commonplace distinction between natural history and natural philosophy or science, and his advocacy of their corporate pursuit prepared the ground for the Royal Society. According to Bacon, natural history is the orderly, but strictly preparatory collection of data, covering as wide a range of experience as possible. Without such preparation the mind will rush into premature generalization on the basis of a few striking or accidentally selected phenomena, like Gilbert, who founded his whole physics on his observations of the magnet. A wide experience of a specific occurrence, such as heat, condensation, the winds and so forth, will include methodologically

significant, if often perfectly ordinary observations. These, duly selected and considered, will suggest generalizations and explanations which can be tested by experiments selected on the same principles. As our hypotheses or ‘axioms’ become wider, penetrating to ‘what is prior and better known in the order of nature’, they become more certain: the method establishes ‘progressive stages of certainty’. This ‘new and certain path’ derives ‘axioms from the senses and particulars, rising by a gradual and unbroken ascent, so that it arrives at the most general axioms last of all’. The aim is ‘not pretty and probable conjectures, but certain and demonstrable knowledge’. The mind ‘must be guided at every step; and the business be done as if by machinery’.¹²⁸

Descartes possibly borrowed from, but also differed from Bacon, who had no recourse to a theory of innate physical knowledge and whose method had no intended application to mathematics. Nevertheless the difference is not straightforward and (as recent commentators have remarked) the notion that the one is the paradigmatic rationalist and the other the archetypal empiricist is not always easy to sustain. Both philosophers disparage the uninterpreted deliverances of the senses, and Descartes as well as Bacon can allow value to well-selected experiment and wide-ranging observations. Moreover, Bacon’s route to a broadly mechanistic view of the world is not obviously consonant with his official programme for gradual discovery, and his confident optimism was hardly less than Descartes’. Locke evidently shared the opinion of many of his contemporaries as to the good sense of Bacon’s principles of experimentation, and in a passage already cited he chided rash generalization in thoroughly Baconian terms:

my Meaning is that we should *not take up any* [hypothesis] *too hastily*, (which the Mind, that would always penetrate into the Causes of Things, and have Principles to rest on, is very apt to do,) till we have very well examined Particulars, and made several Experiments, in that thing which we would explain by our Hypothesis, and see whether it will agree to them all; whether our Principles will carry us quite through, and not be as inconsistent with one Phaenomenon of Nature, as they seem to accommodate, and explain another.¹²⁹

Yet the passage ends with a pessimism unrelieved by Bacon’s salvationist message, and the same pessimism lies behind the sharp distinction between observables and unobservables drawn in the later discussion of probability. There is no talk of Baconian ‘machinery’ to ensure a smooth progression from the observable to the unobservable, no rule for constructing a test of any and every ‘doubtful conjecture’. Locke, it seems, could see the point of a programme of directed experiment, but saw no reason to expect its universal success.

Bacon advocated his method above all for its practical usefulness, the fruits by which it would be distinguished from the barren theories of antiquity. Instead of the few chance discoveries of the past made and developed by practical men—

printing, gunpowder, the compass—he repeatedly promised ‘axioms rightly discovered’ such as will ‘supply practice with its instruments, not one by one, but in clusters, and draw after them trains and troops of works’.¹³⁰ In the *Essay* it looks as if the Baconian argument is subtly scrambled. Gunpowder gives place to quinine in the exemplary trinity of discoveries, presumably because of its less obvious efficacy at saving people ‘from the grave’. More importantly, Locke’s own trinity is actually employed to illustrate the value of the ‘rightly directed’ study of nature, not the haphazard course of that study in the past. Again, he may seem to have intended no more than praise for Bacon’s method, the admired rule of at least some natural philosophers in England, in preferring to ‘Hypotheses laid down at Pleasure’ the work of ‘those, who in this latter Age have taken another Course, and have trod out to us, though not an easier way to learned Ignorance, yet a surer way to profitable Knowledge’. Yet the surely deliberate employment of quasi-Baconian imagery in order to express a point quite opposed to Bacon’s own counts against that simple reading:

In the Knowledge of Bodies, we must be content to glean, what we can, from particular Experiments: since we cannot from a Discovery of their real Essences, grasp at a time whole Sheaves; and in bundles, comprehend the Nature and Properties of whole species together.¹³¹

Despite Locke’s account of the probability of hypotheses, it does not seem that his qualified approval of corpuscularianism in general is founded even on weak arguments from analogy, still less on a Baconian hierarchy of analogical, experimentally confirmed arguments. It is rather that the hypothesis is of its very nature a good one: it goes ‘farthest in an intelligible Explication of the Qualities of Bodies’; and Locke feared that ‘the weakness of humane Understanding is scarce able to substitute’ a better.¹³² Bacon himself recognized the intrinsic virtues of clocks and clockwork as a model: ‘Their wheels seem to imitate the celestial orbs, and their alternating and orderly motion, the pulse of animals: and yet all this depends on one or two axioms of nature.’¹³³ It was perhaps above all its general potentiality for explaining variety and change by ‘one or two axioms’ which accounted for the attraction of corpuscularian mechanism. Its proponents needed no specific point of analogy between existing machines and, say, the blooming of a flower or the reactions of an animal to convince them that the latter could be, or even must be, mechanical.

Another chief element in the attractiveness of corpuscularianism was the possibility that the principles of mechanics, the laws governing matter in motion, are self-evident and necessary, standing to reason in their own right like the axioms of geometry. It is perhaps that conception which should be regarded as definitive of the purest form of philosophical ‘mechanism’, and it promised a brisker, less laborious route to certain and scientific knowledge than Bacon’s officially experimental path. Locke’s response, in effect, was to grant that the conception constitutes the ideal of intelligibility which the true account must

fulfil, but to deny that any available specific version of the corpuscularian hypothesis did more than approach the ideal.

A priori and experimental arguments, however, as even Descartes' famous separation of a piece of wax from its accidents might suggest, seem not always very clearly distinguished in seventeenth-century thinking about method. Despite all his reservations about analysis, Locke in his most corpuscularian moods was prepared to use essentially the same argument as Descartes'. The primary qualities, solidity, extension, figure and mobility, are

utterly inseparable from the Body, in what estate soever it be; such as in all the alterations and changes it suffers, all the force can be used upon it, it constantly keeps; and such as Sense constantly finds in every particle of Matter, which has bulk enough to be perceived, and the Mind finds inseparable from every particle of Matter, though less than to make it self singly be perceived by our Senses.¹³⁴

Other qualities are reducible to these, which by human lights are essential to body. Mechanical impulse is 'the only way which we can conceive Bodies operate in'.¹³⁵ Yet it was clearly Locke's view that such argument only goes so far, for the resultant hypothesis leaves too much unexplained. In the *Essay* the chief difficulties presented are those of how particles cohere, internally as well as with one another, and the unintelligible action by which matter causes ideas in us, especially ideas of secondary qualities.¹³⁶

This appearance of blowing both hot and cold for Boyle's corpuscularianism and standard arguments for it can at first be puzzling, but Locke's position is not really contradictory. He held that the corpuscularian hypothesis was the best available theory and that the arguments for it reveal its virtues; but that they demonstrate neither its capacity to explain everything nor its truth. His pessimism about science can seem unprogressive and old fashioned, but it should be remembered that the kind of test for specific hypotheses of which he apparently despaired was not really achieved in chemistry until very much later. His doctrines represent an effective critical assessment of what had been achieved to date rather than a visionary programme for the future. Quite understandably, he simply could not envisage experiments which might conclusively confirm or refute speculative corpuscularian explanations. He could not even see a way towards a systematic classification of chemical substances, free from the element of mere convenience. The gap between observation and theory seemed to him not only unbridged, but virtually unbridgeable except by plausible speculation. If all this constitutes a certain lack of foresight, yet in the face of existing dogmatism and an extreme underestimate by many (especially the Cartesians) of the obstacles in the path of fundamental theory, such a cautious assessment as his, too sympathetic to be merely sceptical and unconstructive, played an important role in the intellectual progress of mankind. It helped not only to sweep away

‘rubbish’, but to inject into the interpretation of Newton’s admired theory, at its inception, a healthy dose of scepticism and self-criticism.¹³⁷

The topic of testimony, human and divine, led into Locke’s epistemology of religion, broached in [chapter 13](#), above. He offered some appropriately caustic comment on the ‘inverted Rule of Probability’ by which an opinion gets more credit the more hands it has passed through, but the most interesting point relating to human testimony concerns miracles. Stories of miracles are extreme examples of the conflict between testimony and common experience, but Locke in effect argued that, just because the ‘strangeness’ of the alleged facts can be understood as necessary to God’s aim ‘to procure Belief’, it does not, as it normally would, detract from otherwise ‘fair Testimony’. The point might seem a prime example of a theist’s simply assuming what would appear to us to be at issue, that there is a God with appropriate powers and purposes. It might seem only to have met with its just deserts when Hume came to argue that an event contrary to the laws of nature would have to be *so* strange that the good reputation of any number of witnesses could not possibly avail against the weight of common experience (which we can take to include, as Hume himself did not or could not, all future observations and attempts to repeat the circumstances by way of experiment). Hume’s engaging conclusion is that, while a miracle is not intrinsically impossible, we could never reasonably believe the report that one had occurred.¹³⁸ A stronger objection than Hume’s obviously exists: to describe something as contrary to the laws of nature is to describe it as impossible, so that the greater the weight of testimony that a strange event occurred, the greater the evidence that an event so described would not be a miracle. In other words the question arises how Locke, whose epistemology is founded on the assumption of a natural order, could envisage miracles at all.

The answer is not that he uncritically accepted that a God powerful enough to create nature must also have the power to contravene its fundamental laws. His conception of such a law was, as it has just been said, rather purely ‘mechanist’ or quasi-geometrical. To suppose that matter exists which does not obey the laws flowing from its very being would be like supposing that a triangle exists the angles of which are not equal to two right angles. Such mysteries are sometimes ascribed to God’s power, but Locke showed little inclination to do so. Thus where ‘there are certain Relations, Habitudes, and Connexions... visibly included in the nature of the ideas themselves’, then ‘we cannot conceive them separable from them by any Power whatsoever. And in these only, we are capable of certain and universal knowledge’.¹³⁹ Yet every universal truth, as we have seen, is on Locke’s account necessary and in principle knowable, even if not by us. What is ‘visibly’ necessary differs from what is otherwise necessary only epistemically, not ontologically. Locke’s God, as it seems from other writings, interferes with his creation by strictly natural means: even the Flood may be supposed to have been caused by ‘God’s altering the centre of gravity in the earth for a time (a thing as intelligible as gravity itself, which perhaps a little variation of Causes unknown to us would produce).’¹⁴⁰ The logic of this position

requires that the 'variation of causes' itself had natural causes, and so on back to the creation, so that the natural world and its history stands as a seamless whole embodying all God's general and particular purposes from its beginning. Such a view has been held, and was indeed advocated by Locke himself,¹⁴¹ on the grounds that it is more fitting to the dignity of God; but Locke's chief motive seems to have been his philosophical conception of natural necessity.

The only formal definition which Locke ever gave of a miracle is strikingly idiosyncratic: 'A miracle then I take to be a sensible operation, which, being above the comprehension of the spectator, and in his opinion contrary to the established course of nature, is taken by him to be divine.'¹⁴² The motives for this curious definition are revealed in Locke's response to two possible objections to it. The first objection is that it will make what is miraculous relative to the observer, to which Locke replied that any attempt to anchor the miraculous to that which is objectively and in fact contrary to 'the established, steady laws of causes and effects' will be worthless, since we 'have by no means ability precisely to determine what is or is not above the force of any created being'.¹⁴³ The second imagined objection is that such a definition can include 'operations that have nothing extraordinary or supernatural in them'. Locke's reply was that an event is miraculous if in the circumstances it serves as a credential for the bearer of a revelation from God. If a storm abates on Christ's command, its doing so could perfectly well be mere coincidence. Yet enough convenient coincidences would convince most observers that they were none of them accidental, but that each had its role in the divine plan. In the story of Aaron's rod, the achievements of the Egyptian sorcerers (presumably natural, since Locke did not believe in magic) were just as surprising as those of Moses, yet 'when Moses's serpent eat up theirs, when he produced lice, which they could not, the decision was easy'.¹⁴⁴ All this very strongly suggests that Locke wanted to make the concept of the miraculous compatible with two philosophical principles: first, with his central doctrine that sense-experience does not give us knowledge of the fundamental principles of nature; and second, with universal natural determinism or, more precisely, 'mechanism'.

The *Discourse of Miracles* was published posthumously, and in the *Essay* Locke presented a more orthodox face, allowing for 'supernatural Events' which are 'suitable to ends aim'd at by him, who has the Power to change the course of Nature'. Nevertheless it is its being 'contrary to ordinary Observation', together with the suitability to divine ends, which makes the miracle.¹⁴⁵ As in the case of the Flood, God may 'change the course of nature' by natural means. Such expressions as 'contrary to nature' or 'natural' had, for the New Philosophy, a relative as well as an absolute sense. Not everything which happens in a man or horse is in accordance with its nature as a man or horse, but everything is in accordance with the fundamental laws of physics. In Descartes' mechanistic analogy, 'considering the use for which a clock has been intended by its maker, I may say that it deviates from the law of its nature when it does not tell the time correctly'.¹⁴⁶ Consequently the argument in the *Essay* is at least capable of

bearing an interpretation which makes it compatible with that of the *Discourse of Miracles*: which is not, of course, to allow it much force.

The topic of divine testimony is, perhaps surprisingly, of considerable significance for the study of Locke's epistemology, but it may be helpful to preface discussion of it by mentioning some of Locke's own religious principles and purposes. He believed that the existence of God, a wise and benevolent creator, may be deduced from the existence of natural things: above all, things capable of thought. God has the power and right to command, reward and punish his rational creatures, from which springs the conception of a rational and binding morality. There is an afterlife, the Bible is the inspired word of God and Christ is our Saviour. These are the important, central beliefs of Christianity, and their importance chiefly lies, so Locke seems to have felt, in their relationship with morality and their practical efficacy. He seems to have been more interested in works than in faith for its own sake. Whether directly, by rational demonstration, or indirectly, by reflection on the Bible and Christ's teaching, we can be sufficiently acquainted with our duty on earth, and sufficiently motivated to perform it by the consideration that our treatment in the world to come will depend on how far we do so. Beyond these questions, God's purposes are inscrutable, and not our business.

The content of Locke's religious position is closely connected with two of his major concerns, as they are with each other. One is hermeneutics, or the methodology of biblical interpretation, and the other his concern with religious toleration. Locke held that, although the Bible was originally inspired, we should not forget that it was written by men with the general knowledge, beliefs and customs of their time; and that we should employ common sense and reason in our efforts to determine what their words meant when they were written. He did not believe in an inspired tradition of authoritative interpretation, nor, of course, in the authority of the Pope in the matters of doctrine. He took it that, when there are disputes about the content and interpretation of God's word, careful and rational consideration will generally reveal that at least one improbable construction has been placed on it. But in any case other interpretations than one's own should be tolerated, and the issue left available for rational resolution. Only the atheist and the politically dangerous papist deserve to be the object of intolerance.

His evident purpose in the chapters 'Of Faith and Reason' and 'Of Enthusiasm' was to clip the wings of revelation by subordinating it to 'reason', i.e. to the natural faculties in general. First, although in principle any truth, even a proposition in Euclid, might be revealed, it is less certain as revelation than when its 'Certainty is built upon the clear Perception of the Agreement of Disagreement of our *Ideas*'. Accordingly, when a purported revelation conflicts with such natural certainty, it must lose its claim to be revelation. The reason for both these principles is the same: 'the Evidence, *First*, That we deceive not our selves in ascribing it to GOD; *Secondly*, That we understand it right, can never be so great, as the Evidence of our own intuitive Knowledge'.¹⁴⁷ The

nearest possible thing to a conflict between revelation and reason is when a truth is revealed which, employing our natural faculties alone, we would have judged to be improbable. Yet even this is no more than an appearance of conflict. It is rather that one reasonable ground of assent outweighs another, since 'it still belongs to *Reason*, to judge of the Truth of its being a Revelation, and of the Signification of the Words, wherein it is delivered'.¹⁴⁸

It may sometimes seem that Locke puts revelation on the same level as human witness, since trust in either is a matter of judgement. God, however, cannot lie or make mistakes. The word of God is transmitted through fallible and not always very intelligible human beings, but is necessarily honest and draws on omniscience and so is immune to certain kinds of doubt or merely piecemeal acceptance. Sometimes this feature of revelation is emphasized with misleading rhetoric. We are told that, because it is the testimony 'of such a one, as cannot deceive, nor be deceived', it carries 'Assurance beyond doubt, Evidence beyond Exception'. Faith 'leaves no manner of room for Doubt or Hesitation'. Yet what is given so effusively to the 'evidence' of revelation is at once taken away. The sting is in the tail of the argument. Because 'we must be sure, that it be a divine Revelation, and that we understand it right', all this security is rapidly reduced to probability. The term 'evidence' is here in effect being abused, since it stands for a merely conditional certainty. Locke is really arguing that nothing is, in his usual sense, 'evident' about revelation, apart from the platitude that what God has revealed is true.¹⁴⁹ If the *Essay* does not tell us why we should believe that the Bible is inspired, that is no doubt because Locke was not writing against atheists, but against those who promoted or uncritically accepted some arbitrary interpretation of scripture. His target in these chapters is probably diffuse, ranging from Papal infallibility to Protestant prophesy, but somewhere near the centre must stand, not only Quakers, but also those Puritan philosophers, such as Joseph Mede and Samuel Hartlib, who, on the basis of supposedly inspired interpretations of scripture, predicted the millennium and saw its preconditions being fulfilled in their own time. Their explicit claim was that revelation is a third and independently authoritative source of knowledge, standing beside, or rather above, the intellect and the senses. It is received as an unmistakable illumination, and without it, Mede argued, there is no way out of philosophical scepticism.¹⁵⁰ A philosopher as different from Mede as Malebranche later made the not dissimilar claim (but without the pretension to personal illumination) that revelation is the only sound basis for our belief in the existence of material things. Berkeley later spoke of 'the enthusiasm of Malebranche', and it is just conceivable that Locke too had him in mind.¹⁵¹ What is quite clear is Locke's rebuttal of any view that revelation has independent authority, and his rejection of the method of inspired hermeneutics.

Despite this vigorous endorsement of the natural faculties, however, there might seem to be something circular in Locke's epistemology with respect to the role of God. For he often seems willing to imply that the faculties owe their authority precisely to their bestowal by our creator. Even the fundamental

epistemological doctrine that all simple ideas are ‘true’ and ‘real’ is presented in terms of a wise God’s purposes for us, ‘having set them as Marks of Distinction in Things, whereby we may be able to discern one Thing from another’.¹⁵² Again, we are told that ‘God has given me assurance enough of the Existence of Things without me.’¹⁵³ Even the plea for the natural faculties in ‘Of Faith and Reason’ is made in God’s name: to accept that revelation may contradict reason ‘must overturn all the Principles and Foundations of Knowledge he has given us; render all our Faculties useless; wholly destroy the most excellent Part of his Workmanship, our Understandings’.¹⁵⁴ The question arises whether Locke is not in the end guilty of something like the famous Cartesian circle, employing reason to justify reason: by means of my faculties, I can prove the existence of God; since God is not a deceiver, I can trust my faculties.

One critic finds Locke guilty of something even worse than a circular argument. For Richard Ashcraft, Locke’s religious belief underlies and vitiates his entire epistemology, which is tailored and trimmed to fit the presupposed theological thesis that God has made us capable of such moral and religious knowledge as is good for us, but that our cognitive relationship to his natural creation can only be an Old-Testament wonderment at the incomprehensible power of which human observation affords but a glimpse. Locke’s philosophy is represented as above all a conservative response to the forces of change, an attempt to bridge and redirect them so as to confirm in a new way ‘the old foundations of knowledge and certainty’ in ethics and religion. All his arguments, Ashcraft argues, are confined within the presuppositions of faith, in particular, the assumption that our very ignorance is ignorance of an intelligible, divinely created, universal order of things. The conclusion is that a thinker who so naïvely fences himself off from ‘the abyss of scepticism’ and other embarrassments can hardly be taken seriously as a critical philosopher.¹⁵⁵

It is no doubt true enough that the picture of a being whose ‘weak and narrow thoughts’ are indefinitely inferior to the omniscience of God and even to the ‘enlarged understandings’ of angels, and yet which is equipped with all that is really important to it, was one which Locke felt to be both edifying and reassuring. Yet as an attempt to undermine the more usual view of Locke’s realism and rationalism, Ashcraft’s thesis surely fails. It certainly exaggerates the work done in Locke’s epistemological arguments by the references to God, although it is true that those arguments hang together with his assumption of universal causality. What for Descartes is a principle employed as a premise in a formal argument, is for Locke the informal and conventional rhetoric of one Christian writing for, and against, others. The ascription of our faculties to God in ‘Of Faith and Reason’ made an obvious debating point: the fideist and the enthusiast, for all their religious fervour, derogate from the honour and the dignity of the Creator. But, much worse, Ashcraft seems completely to have missed the hard line of Locke’s argument. He quotes Locke’s rhetorical characterization of reason as ‘natural *Revelation*, whereby the eternal Father of Light, and Fountain of all Knowledge communicates to Mankind that portion of

Truth, which he has laid within the reach of their natural Faculties'. Yet he ignores the immediate and more significant sequel: '*Revelation* is natural *Reason* enlarged by a new set of Discoveries communicated by GOD immediately, which *Reason* vouches the Truth of, by the Testimony and Proofs it gives, that they come from GOD.'¹⁵⁶

Finally, and perhaps most interestingly for the secular philosopher, Locke's discussion of faith possesses a striking feature with a bearing on the rest of his epistemology. For in the case of 'enthusiasm' he recognized both the possibility of an illusion of 'evidence' and the need to explain it. What he did, and indeed had to do, was to reduce the purported evidence to mere conviction, and to demand grounds for the conviction: 'For all the Light they speak of is but a strong, though ungrounded persuasion of their own Minds that it is a Truth.'¹⁵⁷ Their persuasions 'are right, only because they are strong in them. For, when what they say is strip'd of the Metaphor of seeing and feeling, this is all it amounts to.'¹⁵⁸ Yet Locke himself can be accused of relying on just the same unenlightening metaphors for his explanation of intuitive knowledge, which 'is irresistible, and like the bright Sun-Shine, forces it self immediately to be perceived'.¹⁵⁹ What could he reply to the sceptical accusation, brought by Hobbes against Descartes, that intuitionists too are only 'sure because they are sure'?¹⁶⁰ That will be one of the questions to be considered in the next section.

Reflections on the definition of knowledge

Locke's approach to knowledge and belief might appear to differ from present-day approaches in such fundamental respects that it might seem lost labour to engage in any detailed comparisons. One might suppose that the task of the philosophical critic is to denounce Locke's intuitionist, infallibilist and foundationist blunders, praise him a little, perhaps, for his recognition of the importance of probability and his general opposition to dogmatism, and pass on. Yet it could be that we have most to gain precisely from the effort to do justice to what at first appear his least tenable principles. Philosophical error is often implicated with insight, and overreaction to error can hide important truths from us, sometimes for generations. Typically, our reading of past theory becomes mere parody. The present section will comprise a general explanation of how such a thing might have happened in the theory of knowledge, and the argument will be developed in the succeeding sections on perceptual and universal knowledge.

An obvious starting-point, perhaps, is the failure of fit between Locke's conceptions of knowledge and belief and the ordinary conceptions to which, one might suppose, even a philosopher needs to assign at least provisional authority. For Locke, as we have seen, knowledge is much more confined than what we ordinarily count as knowledge, and belief is much narrower than what we ordinarily count as belief. Unlike Locke, most of us who have any ideas about the subject at all suppose ourselves to know, and not merely to believe, that a successful general called Julius Caesar ruled Rome and was assassinated more than two thousand years ago. Moreover knowledge and belief are not, as we normally think of them, mutually exclusive categories. Belief is, rather, a condition of knowledge. Recent systematic philosophy of knowledge and belief mostly takes that relationship for granted: knowledge exists when true belief arises in certain specific circumstances. The problem is to specify the circumstances.

It might be argued that these differences are superficial, or at least easily explicable. If Locke's usage of 'belief' comes close to what is still ordinary usage of 'opinion' when we speak, for example, of something's being a 'matter of opinion', then perhaps his principle that what we know we do not believe is

not so paradoxical. More importantly, if it is one of the necessary conditions of someone's having knowledge that his or her belief (in the now standard sense of 'belief') be justified or well-grounded, then we can readily understand, without supposing that Locke employed the term 'knowledge' in any but the ordinary sense, how and why he should have come to hold that much of what is ordinarily counted as knowledge is not really knowledge. The notion of justification is evaluative or normative, so that standards for what is 'justified' may vary without the notion varying. Locke, it might be said, simply demanded that belief met higher standards of justification than the normal ones before he was prepared to do it the honour of calling it 'knowledge'. Locke's scepticism was not extreme, but on this view even an extreme sceptic, someone who asserts that we know nothing, or nothing but that we are now having certain thoughts and experiences, is using the word 'know' in the ordinary sense. Just because 'know' is evaluative, we cannot refute the sceptic by appealing to the ordinary usage of 'know'.

That now familiar point deserves to be taken, but it neither sets out the reason for Locke's dichotomy between knowledge and belief nor intimates what we can learn from it. He was not interested simply in denigrating some beliefs as 'probable' which we would normally treat as 'certain'. As we have seen, he insisted that for all practical purposes high probabilities can and do serve as well as certainties. He was interested in a difference, not of degree, but of kind. What is now difficult is to achieve a sympathetic understanding of his account of that difference. One obstacle, of course, lies with the assumptions about logical form which are embodied in his definitions of knowledge and belief. The claim that knowledge consists in the 'perception', and belief in the 'presumption' of a relation between ideas is not obviously helpful to anyone reluctant to make concessions to Locke's subject-predicate model for wordless mental 'affirmation'. Even in its own terms, the model seems stretched past breaking-point in its representation of 'sensitive knowledge' as the perception of a relation between the idea of an object and the idea of existence. Nevertheless the definitions point clearly enough to Locke's motivation. Very roughly, the basic thought is that we have knowledge when and only when it is appropriate to speak of 'perceiving' the truth; whether when we literally see (feel, smell, etc.) that something is in fact so, or when we metaphorically 'see' that something is necessarily so, as in mathematics or logical reasoning.¹⁶¹

In knowledge, in other words, the truth or fact or state of affairs in question is itself perspicuously presented to us by our faculties. Opinion is less logically self-contained. In 'judgement', however successful, we get at the fact in question more circuitously, through grounds extrinsic to the fact itself: 'in all the parts of Knowledge, there is intuition; each intermediate *Idea*, each step has its visible and certain connection; in belief not so. That which makes me believe, is something extraneous to the thing I believe.'¹⁶² Another word Locke used for the perspicuity of knowledge is 'evidence', although (as we have seen) this term occasionally spilt over into his discussion of 'apparent probability'. It is significant that it did so (as we shall see), but in general the evaluation of

Locke's dichotomy of knowledge and belief has to be an evaluation of the metaphors of sight, light, evidence and the like as they are applied to knowledge other than perceptual knowledge or, indeed, (since perception and knowledge are *never* quite the same) to knowledge generally.

The denunciation of the metaphor of sight for knowledge is something of a war-cry among 'anti-foundationists', for whom the very term 'epistemology' is commonly reserved for a wrong-headed search for the absolutely dependable, evident foundations upon which all other knowledge is or might be built. This search was supposedly initiated by Descartes and only rather recently brought to an end. For many, it has been replaced by the view that knowledge is a system or web the parts of which may prop one another up, but none of which is intrinsically fundamental or beyond question. If we *treat* some elements of the system or its general structure as fundamental, unquestionable or necessary, then that is a matter of decision or tradition or utility or ideology or our way of life, rather than a matter of their intrinsic obviousness to sense or to the eye of reason. The view which will be advocated in the present work is that knowledge does indeed have foundations, although not unquestionable or infallible ones. Crudely, there are epistemically primitive deliverances of our cognitive faculties which do, as such, have a kind of perspicuity or 'evidence' to us; but our faculties are not infallible. Without such defeasible foundations, knowledge would be impossible. The argument is not an attempt to revive 'epistemology', since that branch of philosophy has not ceased to thrive. It is an unhelpful and by now rather silly piece of rhetoric to pretend that anti-foundationist epistemology is not epistemology. Even if all forms of foundationism were false, philosophers would still need to understand what knowledge is, and the conditions under which it arises. As it happens, reflection on past foundationist theory (which, so far from beginning with Descartes, goes back to the ancients) is not useless for the purpose.

A standard objection to the notion of infallible faculties is that it is a mere sideways shuffle in the face of scepticism. The problem of scepticism simply returns as the problem of how we are to know that the infallible faculty really is operating on those occasions when it seems to us to be operating. The argument mirrors Locke's own argument against the claims of revelation, except that it rejects Locke's implicit contrast between the 'light from heaven' and the 'light of nature'. What is evident or revealed is doubtless true, but how is the 'evidence' of *either* revelation *or* intuition to be distinguished from mere conviction? In both cases, it seems, those who appeal to the inner light and nothing more can be accused of being 'sure because they are sure'. Intuitionism, like the doctrine of revelation, seems to postulate an impossible cognitive state, a kind of self-validating conviction, which is at the same time useless in that another cognitive state, deluded opinion, can be confused with it. The fault may seem to lie with the conception of propositions as psychological entities rather than logical ones. The modern separation of logic from psychology seems to present a choice: by 'evidence' we must mean *either* a property of some cognitive states *or* a logical

feature of some propositions, but we cannot mean both at the same time. If logical 'self-evidence' is indeed a sure criterion of truth, its presence is not automatically given to us as a phenomenal feature of our thinking, but has to be discerned in the propositional object of our thought (or, on some conventionalist views, created by *fiat*).

Such objections to intuitionism have contributed massively to the presuppositions of twentieth-century epistemology and philosophy of mind, to the extent that Locke's theory is liable to seem not only alien, but foolish. An epistemology apparently grounded so directly on the tautology, 'If I perceive it, it must be true', may seem to us no better than a primitive blunder. Yet an effect of the modern argument has been to cripple our understanding of knowledge. For the natural, if not inescapable metaphor of sight does after all have considerable point. Before considering what that point is, it may be helpful first to take a look at some recent epistemology. There is space only for a commentary conducted in terms of the utmost brevity on the continuing and mountainous debate as to the necessary and sufficient conditions for knowledge.

The argument took a fresh turn in 1963 with a very simple question. Since we can be justified in believing what turns out to be false, and since it is possible for true conclusions to be drawn justifiably from false premises, it is clear that there is more to knowledge than justified true belief. What more?¹⁶³ A possible response to this challenge is to try to tighten up the notion of justification by arguing that a belief is not in the requisite sense 'justified' if it is not adequately grounded on *true* premises or presuppositions, by true principles: in other words (employing a looser, less technical sense of 'justified'), that it is necessary only to add the condition that the justified true belief does not depend on false lemmas. That amendment, however, is evidently not enough. For, if a belief is to constitute *knowledge*, it must not only be independent of *false* lemmas, but every lemma upon which it rests must be *known*. Yet to include that as a condition of knowledge would, of course, be circular.¹⁶⁴

Here there might seem to have been an opening for an anti-foundationist proposal that a belief constitutes 'knowledge' precisely in virtue of its membership of a coherent web of interdependent beliefs: in effect, that enough beliefs taking in one another's washing count as 'true', 'justified' and 'knowledge' all at once. But a different line of thought turned even many anti-foundationists in a less heroic direction. Here the leading idea is that, for a belief to constitute knowledge, our possessing it must depend causally on the fact or state of affairs in virtue of which it is true. According to one development of this idea, what makes our belief that Julius Caesar was assassinated count as knowledge is not the strength of our reasons for holding it, such as they may be, but its having been appropriately, if indirectly, *caused* by Caesar's assassination. A schoolboy, let us suppose, knows it because he has been taught it by his teacher, the teacher's belief is due to her having read various books, and so on back to the original witnesses' beliefs which were brought about through the normal operation of their senses stimulated by the actual event. At its most ambitious,

this causal approach to knowledge offers not just to supplement, but to supplant the condition that the belief be justified. Knowledge is presented as a straightforwardly natural occurrence, definable in terms of natural, rather than logical or normative relations.¹⁶⁵ Moreover the sceptic may be characterized as someone who fails to grasp this essential characteristic of knowledge, pinning everything on justification. But even apart from such pretensions, some have thought that the ‘causal theory’ best explains the extent of our ordinary conception of knowledge, and the relaxed attitude towards the possession of reasons which may seem to be embodied in it. I can, for example, know something just because I was told it, and can continue to know it even after I have forgotten when and how I learnt it, or why I first believed it.

The thesis that the central and explanatory core of the notion of knowledge is the causal dependence of knowledge on what is known has become more complicated as its proponents have tried to meet a variety of difficulties. One well-known problem arises because factors lying quite outside the ‘causal route’ between fact and belief can determine the question whether the belief constitutes knowledge. For example, if I see clearly a pound coin lying in my hand, then in normal circumstances my belief that I have a pound in my hand does constitute knowledge. But in a particular case in which, unknown to me, there are a large number of sensibly indiscriminable counterfeits in circulation, some even lying in the pocket from which I have just taken the genuine article, it is clear that I do *not* know that I have a pound in my hand. Yet there is no difference between the two cases with respect to the way in which my having a pound in my hand has brought about my true belief that there is a pound in my hand, or, for that matter, with respect to the ground of my belief. The difference is that in the second case there is something too chancy about my having been right.

The same sort of point can be drawn from the possibility of faculties which sometimes work properly but more often erratically. If the mechanism of someone’s colour vision functions in the normal way only spasmodically, on other occasions giving false or randomly true impressions of colours, then perhaps he should be said *never* to have perceptual knowledge of the colours of things. He does sometimes see their colours and (let us suppose) always believes what he sees, yet his colour vision is not to be trusted. Examples like these have encouraged the adoption of a rather wider criterion of knowledge than a pure “causal theory” would propose: in order for a belief to constitute knowledge, it must in some sense be ‘no accident’ that the belief is true. Yet such a criterion at least calls for further explanation and refinement. One might, for example, ‘accidentally’ acquire a true belief as a result of glancing by chance in a certain direction, or because some distant object happened to catch the light of the sun. Such fortuitous knowledge is none the less knowledge.

Another significant objection to a simple causal theory focuses on the problem of explaining the notion of an *appropriate* (or, as it is sometimes called, ‘non-deviant’) causal route from fact to belief. For states of affairs can cause beliefs that such states of affairs obtain by routes which preclude those beliefs from

constituting knowledge. For example, a bricklayer, knocked unconscious by a falling brick, may be caused by the pain to dream of being struck by a falling brick, and consequently wake in a state of confusion with the true belief that such a thing has happened. Why is this belief not knowledge?

In the case of perceptual knowledge, a first answer to the problem may seem simple: a 'non-deviant' causal route is the normal one for the sense in question. A less vulnerable answer, however, has recourse to the general notion of a cognitive faculty. For we can allow that the mechanism of a cognitive faculty may be highly abnormal (e.g. if medically repaired), and yet still be productive of knowledge. There is nothing impossible in principle about an artificial functioning 'eye'. This footnote to the causal theory seems to have rediscovered a platitude which traditional epistemology took for granted, that knowledge is always the product of our cognitive faculties. Indeed, the 'counterfeit' and 'deviant causal route' objections to a simple causal theory have given rise to an attempt like the Stoics' to define knowledge in terms of discriminatory capacities and, in general, the capacity to discriminate truth from falsity in the circumstances. Yet there are cogent lines of objection to this approach too. First, on any ordinary conception of a cognitive faculty or capacity, we would wish to distinguish the possession of the faculty from the satisfaction of the conditions in which the faculty might be exercised. The presence or absence of counterfeits might affect the latter, but only a change in *us* (or at least in the mechanism of the capacity) could constitute a change in any of our discerning capacities. That is not to say that it would be utterly unintelligible to employ a different way of conceiving of a 'faculty', but that, if we do so, the explanatory force of an appeal to the notion becomes problematic. Like the attempt to embody the 'no false lemmas' rule in a strengthened notion of *justification*, the attempt to cover considerations of quite different types under the umbrella of a suitably reshaped notion of a *discriminatory capacity* appears explanatory only in so far as the different nature and significance of those considerations is masked by the conceptual distortion.

Yet there is a more serious difficulty for any 'faculty' approach, regarded as a refined version of the causal theory and so as a *rival* of a 'justification' approach to the problem of the definition of knowledge. A cognitive faculty is both something more and something less than a belief-producing causal mechanism. That it is something less (to take the sense of sight as our example of a cognitive faculty) is illustrated by its being possible to see something without believing our eyes: perception is possible without perceptual belief. That it is something more (again, to take the case of a sense) is illustrated by the role in sense-perception of *sensation*, that presentation of things in consciousness upon which perceptual belief is grounded. (These are topics which will be dealt with more fully in [Part III](#), below.) Thus, an artificial eye might function on very different principles from a natural eye, but it will serve and count as a part of the mechanism of *sight* only in so far as its functioning leads to a stimulation of visual sensations appropriate to the objects to which it responds. Consequently,

even in the relatively simple case of perceptual knowledge, there is no hope that the notion of a 'non-deviant' causal route from fact to belief (via the functioning of a discriminatory faculty) can be adequately explained independently of the notion of the grounding or, in some sense, 'justification' of belief.

The same conclusion can be drawn with respect to inferred knowledge, and 'justification' in a more straightforward sense. The schoolboy who knows that Julius Caesar was assassinated because his teacher tells him so is employing his reasoning faculty, or at any rate manifesting an intelligent grasp of his epistemic situation, in so far as he rationally trusts his teacher, i.e. trusts with good reason. And there is at least some cogency in the view that, unless he believed with reason, he would not have knowledge. For example, if he had believed his teacher not because of any such intelligent trust, but just because the teacher had hypnotized him (and, let us add, had done that just in order to instil into him the truth), the causality of the belief would not have been appropriate either to the belief's being the outcome of the operation of a cognitive faculty or to its being knowledge. The point can be extended to the whole chain of inferences and experiences which lead back to Caesar's assassination. It is not necessary that every individual person involved in the chain should have had the knowledge in question. Scribes may have copied out a history without believing or caring whether what they wrote was true. But some, at least, of their readers must have *reasonably* presumed that what they wrote owed its content to the knowledge, if not of themselves, then of others.

Much the same goes for such inferential knowledge as does not involve testimony. In quick illustration, let us suppose that vaporized mercury is sensibly undetectable and commonly causes paranoia; and that the specific form of paranoia caused consists in the belief that one is being subjected to vaporized mercury. That gives us two ways in which being subjected to mercury vapour can give rise to the belief that we are being subjected to mercury vapour. One is irrational, bypassing our cognitive or discriminatory faculties (and working directly, as Locke might have put it, on the fancy): the other is rational, as we employ (say) our knowledge of the effects of mercury together with our observation that the people immediately around us have become irrationally convinced that they are being subjected to mercury vapour. Only the second can give rise to knowledge.

Thus, despite their unseemly rivalry, the 'causal' (or 'no accident') approach and what might loosely be called the 'rationalist' approach to knowledge each has a need for the other. One way in which their spurious antagonism has been presented has been through a contrast between 'externalist' and 'internalist' theories of knowledge. These terms are almost intrinsically unclear and misleading. From one point of view, for example, the condition that knowledge is consequent upon the exercise of cognitive faculties is 'internal', because faculties are internal to their possessor. But from another point of view it is 'external', in so far as the condition is a function, not of the subjective mental state of the believer, but of the objective causality of that state. Yet one way of characterizing the difference between 'externalist' and 'internalist' theories of

knowledge which is at least reasonably clear is as follows: according to the former, the only 'internal' or 'subjective' condition or criterion of knowledge is *belief*, with more or less conviction; while, according to the latter, the only 'external' or 'objective' condition of knowledge is the *truth* of what is believed. On this understanding, however, there is plenty of space between internalism and externalism. The extremes can be illustrated by, on the one hand, the attempt to define knowledge as justified true belief and, on the other, the simple causal theory. The position adopted in the present argument is that the truth does in fact lie between these extremes; and that, whatever their faults, the classic theories of knowledge constructed before the excesses of the twentieth century lie there too. Certainly that is where Locke's theory lies.

One of the more prominent recent attempts to present a worked out externalist theory of a broadly 'no accident' type is that of Robert Nozick. The core of his proposal is the claim that S knows that *p* under the following conditions:

- (1) *p*
- (2) S believes that *p*.
- (3) If *p* were not true, S would not have believed it.
- (4) S believes that *p* because *p* is true. (Taking that to give the sense of Nozick's somewhat bewildering conditional, 'If *p* were true, S would believe it.')

The last two conditions are united under the notion that a belief constitutes knowledge if it 'tracks the truth' in possible worlds 'near' to our own. In effect, the third condition is explained as stating that if *p* had been *false*, it would not easily have been the case that S *would* have believed it; while the fourth is taken to state that, given that *p* is *true*, it would not easily be the case that S would *not* believe it.

Criticism of this proposal, for example by Christopher Peacocke,¹⁶⁷ has suggested in effect that Nozick's theory faces difficulties analogous to the difficulty faced by a simple causal theory of independently characterizing the class of 'non-deviant' causal routes between fact and belief. For the theory entails quite implausible interpretations of some examples unless we set principled limits on our understanding of which possible worlds are 'near' to our own (i.e. of what would or would not easily be, or have been, the case). Arguably an ineliminable determinant of such limits must be the supposedly jettisoned condition that knowledge be well-grounded. In fact, Nozick helps himself to the notion of 'methods' or 'ways' of 'arriving at' beliefs, a notion needed even to explain why fortuitous knowledge is not excluded by his fourth condition. Typically, when we possess such knowledge that *p*, then, on any straightforward understanding of the situation, we might easily have believed that not-*p*. Nozick's solution is to modify his fourth condition thus: given *p*, then, *on the supposition that we employed the same method of arriving at a belief*, it would not easily have been the case that we believed that not-*p*. Even if the form of this solution seems attractive to us, however, its application to the case of fortuitous

knowledge is unclear. It might well be that, employing the same *method* of looking around for evidence, we might easily have believed that not- p (i.e. if we had happened not to notice what we in fact chanced to notice). So the solution seems to hang on a circular principle of distinguishing 'methods': in effect, 'noticing that p ' has here to be a different 'method' from 'not noticing that p '.

In general, Nozick's notion of a method of arriving at belief departs radically from the natural and normal notion upon which it trades. In dealing with another type of case, that of overdetermined belief, he lumps such causes of belief as prejudice and love in with the acquisition of evidence, counting all as 'methods'. The case is that of a father who truly believes with good evidence that his son is innocent, but who would have believed in his son's innocence even if he had been evidently guilty. It is both more natural and more explanatory simply to say that the father's present belief constitutes knowledge because it is well-grounded, than to say that it does so because, if his son had been guilty, he would only have arrived at a belief in his son's innocence by employing a different method. A 'method', in any ordinary sense of the term in this context, is a cognitive procedure intentionally adopted by an enquirer as a way of arriving at well-grounded beliefs. No one could adopt prejudice as a method.

The fundamental fault of Nozick's approach can be illustrated by the cases previously considered. A schoolteacher who hypnotizes pupils in order to instil the true belief that p arguably fails to instil knowledge that p even though (given her determination, competence, hypnotic ability and upright character) the pupils would not easily have believed that not- p and, had p been false, would not easily have believed that p . A supporter of Nozick would presumably claim that the pupils *would* have knowledge, just because of the reliability of the 'method' involved. He would probably accept that their having knowledge depends in part on the teacher's not being the sort who might easily, on a whim or in another mood, have chosen to instil falsehood. Yet we could equally well argue that another mood would have constituted another 'method', so that even a teacher liable to periods of irresponsibility could deliver knowledge hypnotically. There is no way of arbitrating between such rival judgements without circularity, since being hypnotized and the like is not a 'method' of acquiring knowledge in any ordinary sense. In purging the notion of a method of its usual implications, Nozick has simply made it indeterminate. Related to these considerations is the point that, even if the hypnotically instilled true belief counts as knowledge, it is not a paradigm or central case of knowledge. What that means will shortly be discussed. But any inclination to count the pupil's belief as knowledge derives at least some of its force from the case's being one of the transference of a belief from one person to another. The teacher is assumed to know the facts in a perfectly straightforward way, and some of the status of that knowledge rubs off. That seems to have been why divinely instilled beliefs, innate or inspired, were generally taken to count as knowledge. (Locke, on the other hand, was struck by the difference between such a circuitous route and that operation of natural faculties which brings with it the possession of grounds.) Hence our second

example, in which no such transference of knowledge occurs, falls out rather differently. The paranoid, mercury-induced belief that one is being subjected to mercury vapour might 'track the truth' with the best, but would not constitute knowledge because it is not acquired in the right way. It is neither well-grounded nor, *a fortiori*, the product of a cognitive faculty.¹⁶⁸

Considerations of the 'no accident' type and considerations of rationality or grounding, however, commonly run in parallel. For example, it is said that Mendel's original hypothesis in genetics was based on an inadequate sample of evidence. So Mendel believed, but did not *know*, that Mendel's Law is true. On a Nozickian analysis that is either because it could easily have been the case that Mendel had not believed it (i.e. if the same method of sampling had turned up less fortunately suggestive results); or because, if Mendel's Law had been false, it could easily have been the case that Mendel had still believed it (i.e. a certain range of different laws could easily have supplied Mendel with the same evidence); or, of course, for both reasons. Yet it is difficult to see that these Nozickian explanations are more basic than (not to say, independent of) the normative statistics, or evaluation of Mendel's evidence, upon which they are in fact based.¹⁶⁹

Here Nozick might say that his formulation is explanatory just because it is more general. There are other ways in which it might be the case that a scientist's true belief could easily have been false: for example, even if Mendel's sample had been statistically adequate, Mendel himself might have been such a bad judge of evidence that he might as well have drawn the right answer out of a hat. In that case too he would not have *known* that his conclusions were true. Whether or not Nozick would authorize this imaginary reply on his behalf, it seems to be very much in the spirit of his enterprise, for he believes that we should be seeking a single, simple criterion of knowledge which covers all cases of knowledge. In that way alone can we reveal what he calls the unity of the concept of knowledge. Yet, quite apart from the arguments which demonstrate that in itself Nozick's criterion fails to fit (let alone explain) crucial cases, he has a mistaken conception of the philosopher's task or, more particularly, of what it can be to reveal the unity of some notion. First, it is not necessary, in order to show that knowledge is one thing, to bundle together the various characteristic features of knowledge under a single formula. The natural effect of such an ambition has been to produce distorted accounts favouring one feature over the others: e.g. the consideration that knowledge must have the right origins over the consideration that it must be well-grounded, or *vice versa*. As the consequences of such distortion become apparent (and one thing to be said for the present philosophical atmosphere is that the paradoxical consequences of all but our deeper errors are soon revealed—we live in the age of the counterexample) the single criterion becomes perforce less rich, and its application to the cases so ingeniously complex, that the significant features of knowledge become obscured rather than revealed. Nozick's achievement is to have produced a general criterion of almost superhuman aridity, but its application and

justification, in so far as it is plausibly done, calls back the excluded rivals in disguise. Unless as an object-lesson, it tells us rather little about knowledge.

Its failure suggests a quite different model for philosophical explanation from Nozick's. In order to grasp what *knowledge* is, it is indeed helpful to construct a list of the conditions or, as it might be better to call them, characteristics or features or aspects of knowledge, but there is no virtue in making the list as short as possible—rather the reverse. The unity of the notion is then to be demonstrated by our uncovering as far as possible the inter-relations and interdependences of these aspects without obscuring their differences. The metaphor of an 'aspect' is a good one, for in general how a thing looks from the north, although different from, is at the same time systematically related to, how it looks from the east, west or south.

Applying this model for philosophical explanation, we can say (for a start) that knowledge (1) is belief which is true, (2) is rational, justified or well-grounded, (3) is based on no false lemmas,¹⁷⁰ (4) arises in consequence of the operation of cognitive faculties (each, of course, with a mechanism), and (5) derives in an appropriate way from the state of affairs (if any) towards which it is directed.¹⁷¹ We can add that the faculty must be dependable, and should explain that its being dependable on a particular occasion may lie in the circumstances of its exercise (as in the case of the counterfeit coins) rather than in the state of its own mechanisms. The unity of this list is revealed when the interconnections among its items are revealed, e.g. when it is demonstrated that the consideration that knowledge involves belief with an 'appropriate' derivation can be explained only with reference to the other features of knowledge, to its being characteristically well-grounded, and consequent on the operation of our faculties. But there is nothing to be gained, for example, by constructing a forced and artificial notion of 'justification' which incorporates the condition that a 'justified' belief is based on no false lemmas. Again, a belief which is rationally based will also be a belief which is the exercise of a cognitive faculty, but its being the former and its being the latter are not conceptually the same.

An extremely important feature of such a list of the characteristics or aspects of knowledge is that it is open-ended. In other words, a better grasp of what knowledge is could involve our adding to the list a further item which we should expect in some way to flow naturally from other items, but at the same time to increase our understanding of them. Such an addition will be consonant with, but not reducible to, the other items; as they are consonant with, but not reducible to, one another. The possibility of such extension of the list brings us back to Locke and to what we might learn from the metaphor of knowing as seeing. Yet first it is necessary to identify a further mistake about philosophical explanation which is embodied in Nozick's approach and in that of most other contributors to the recent dispute about knowledge. That mistake is the presumption that, if the concept of knowledge is unitary, the characteristic features of knowledge will be found in each and every case of knowledge.

The thought that it might be wrong, in the case of many predicates, to expect to find necessary and sufficient conditions for their satisfaction has been a familiar part of the philosophical scenery since Wittgenstein suggested that the particulars falling under a concept (his most famous example is *game*) may be related by 'family resemblances'. These are features or points of resemblance characteristic of the class, but none of which is necessarily present in every member of the class. A game need only have *some* game-features, *some* points of resemblance to *some* other games. There can therefore be central or paradigm or typical cases, possessing a high proportion of the 'family resemblances', and less central cases, closer to the borderline, with a correspondingly smaller proportion. Wittgenstein himself constructed his model very much with an eye on cognitive concepts (in particular, on *understanding*), and for some it explains the inconclusiveness of the debate over the necessary and sufficient conditions of knowledge. Nevertheless there is something deeply unsatisfying about Wittgenstein's model as it is generally propounded. The metaphor of a network of criss-crossing resemblances, denser at the centre, might seem to explain how the instances are held together, but it fails to explain the difference between relevant and irrelevant resemblances, and so, in any real way, the unity of the concept. It could have been the case, even if it is not, that as many games were played with balls as are played for enjoyment, yet a game's employing an object of a particular shape is conceptually irrelevant to its being a game as its being played for enjoyment is not. That so many games are ball-games depends presumably on the contingent potentialities for interaction between round, more or less elastic objects and human beings, the psychological effects of certain sorts of such interaction, and so forth. Wittgenstein's failure to meet the sense that there are notional *unities* to be revealed is of course consonant with his poor opinion of the pretensions of philosophy to explain anything. Understandable reaction against such pessimism may help to account for the fervour with which the search for a comprehensive and simple criterion of knowledge has more recently been pursued.

The model here proposed, however, is at least a reinterpretation of the usual 'family resemblances' model, if it is not a different model altogether. So much should already be clear from the importance attributed to the task of revealing interdependences between the various features of knowledge. Moreover, on the presumption that knowledge is a unitary notion, if a feature of one kind of knowledge which seems relevant to its being knowledge does not appear in knowledge of another kind, some explanation is called for. One form that such explanation might take would be the identification of some broadly analogous feature in the second kind of knowledge. To pick on a notorious example, the condition that knowledge must derive from its object cannot be satisfied by *a priori* knowledge. Even if mathematical facts are facts, they are not facts which can impinge causally on the world or, therefore, on us. Nozick claims for his 'truth-tracking' criterion the merit of uniting *a priori* and factual knowledge: someone knows that $7+5=12$ if and only if it is *became* $7+5=12$ that he *believes*

that $7+5=12$. Yet that is an extraordinarily blank assertion if it is made (as it is) with no account of what relationship is expressed by the 'because', or of what in the world makes it the case that the relationship holds. Only such an account, given at a lower and richer level, will enable us to judge whether there is indeed a significant analogy between factual and *a priori* knowledge in this respect. Otherwise the explanation is at best like the abstract 'proof' that there are aesthetic facts because some aesthetic propositions are true, a claim which does nothing to explain either the resemblances or the differences between alleged aesthetic facts and truths and those which are less problematic. Consequently it may be more illuminating to think of some different sorts of knowledge as having analogous features, than as sharing the same feature. In that way, at least, room is left for an exploration of the extent of the analogy.

Nevertheless the pursuit of analogies in the name of unity will not take us very far from the principle which Wittgenstein rejects, that there are defining characteristics of knowledge present together in each and every case of knowledge. The difference between a characteristic's being the same as another and its being analogous to it is indeterminate. The more interesting possibility is that of knowledge which lacks absolutely some characteristic feature or features of knowledge. Even to envisage such a possibility might seem to commit us to Wittgensteinian laxity and disorder, but fortunately there is another way of maintaining discipline and unity. That is by a distinction between primary and secondary knowledge, such that there could be no knowledge at all without primary knowledge.

Consider again the example of *games*. Playing (whatever its biological or social origins and functions) is something characteristically and of its essence done for its intrinsic enjoyment, for its own sake and as an end in itself. Primary cases of games (primary games) are played in that way. Rules, competition (with winning and losing) and even gambling enhance, rather than detract from this feature of primary games, although they are not necessary for it. At the same time they supply the conditions for the secondary playing of games: for the purpose of training soldiers or diplomats, for earning a living, for deciding matters of life and death or the like. (Conversely one can understand how war, or speculative buying and selling, is for some just a game.) To understand what a game is, or the unity of the notion of a game, one had better understand primary games first. Not only would the profession of the gambler or sportsman for whom playing is just work be impossible unless some people saw gambling as a pleasure and sport as an end in itself, but it would be unintelligible why a rule-governed procedure, employed, for example, in testing certain skills or for deciding whose life is to be saved, should be properly describable as a 'game', unless that procedure, or something sufficiently like it, could be conducted in other circumstances as a primary game for its own sake. So the distinction between primary and secondary games is an important one if we are to discern the order and unity which exists among actual and possible games.

Is it as useful to draw a distinction between primary and secondary knowledge? One might think of doing so on the basis of the difference between immediate and inferred knowledge. Yet although it is true (*pace* the anti-foundationists) that, unless there were immediate knowledge, there could be no inferred knowledge, the dependence of inferred knowledge on immediate knowledge does not have as a corollary the absolute lack of some central characteristic of knowledge. A more promising candidate for the status of secondary knowledge is perhaps innate knowledge (whether or not there actually is such a thing) which *ex hypothesi* has arisen otherwise than through the operation of our cognitive faculties. By the same token it lacks subjective grounding, a point which is the more important for the present purpose just because it could be argued that the *causality* of innate truths, whether we owed them to a benevolent divinity or to natural selection, would be sufficiently analogous to the exercise of our faculties for it to be ‘no accident’ that we believe as we do. Here we come close to Locke’s reasons for rejecting the contemporary doctrine of innate knowledge. In being subjectively ungrounded a merely innate belief must lack the perspicuity and ‘evidence’ required to satisfy his definition of knowledge as ‘perception’. If, on the other hand, a belief truly possesses the evidence which his opponents generally claimed for their innate principles, then, as Locke argued, the hypothesis of innateness is redundant and the belief can be ascribed to a faculty of reason. In being evident, the beliefs do not strike us as innate, but as rational.

The notion of the ‘evidence’ of knowledge is at least the notion that, when we know something, we know how we know it. The seeming circularity is important: ‘evidence’ is not to be explained simply in terms of a second-order belief about the origins of the first-order belief. Rather, the origin and ground of our belief is in a certain sense perspicuous to us. There must be nothing of a hunch about either the first-order or the second-order beliefs involved. For knowledge that *p* to be perspicuous it is not enough that its possessor should have knowledge of how it was acquired, since that knowledge too must be perspicuous or primary knowledge. The circularity shows that a reductive explanation of ‘evidence’, and so of primary knowledge, is impossible. Nevertheless, something can be said in its explanation: for a start, that it is characteristic of the exercise of our senses, memory and intelligence. When I believe my eyes, there is no mystery about the source and basis of my belief, and the same is true when I perceive why something is necessarily so, e.g. when I *understand* the truth that, if A is larger than B, and B than C, then A is larger than C. Locke simply refused to accept that anything lacking this central feature of knowledge could count as knowledge. Even ‘habitual knowledge’ he restricted to cases in which its possessor ‘can remember that he once knew it’, so that it has in effect a *present* grounding: ‘he knows it in a different way, from what he did before’.¹⁷²

A distinction between primary and secondary knowledge, however, can preserve the point of Locke’s respect for this perspicuity and at the same time

allow a more natural and relaxed view of what can constitute knowledge. A memory-belief stripped of all recollection of its acquisition may contribute as much as any memory to our successful functioning in the world. Nevertheless it is second-class memory, secondary knowledge in so far as it requires for its status as knowledge the context of other more full-blooded recollections. In memory, Locke tells us, the mind revives perceptions 'with this additional perception, that it has had them before'. It may seem easy to refute him with examples of unconscious memory and the like, and so to prove that the essence of memory lies in certain appropriate effects of past on present through the medium of a cognitive mechanism, rather than in the subjective presentation of the past as such. Yet if we lacked *all* perspicuous remembrance of the past, we would have no knowledge of the past at all, even if some of our knowledge of the world was acquired in the past. For we would lack even a conception of the past. We need to read Locke as giving an account of primary memory-knowledge and, more generally, as allowing into his category of *knowledge* nothing but primary knowledge.

Perhaps a clearer illustration of the distinction between primary and secondary knowledge is supplied by a phenomenon which has recently received discussion by psychologists and philosophers of perception. 'Blindsight' occurs when a certain sort of brain-damage has restricted its victim's field of conscious vision. The subject at first supposes that he cannot see what lies in the occluded area, but if pressed to locate or identify an object by what he initially takes to be guesswork, then he can do so with considerable success. We can therefore say that the object is 'seen' but not registered at the level of conscious sensation. Such unconscious rapport with the world, acquired through the familiar mechanism of sight (or a large part of it), makes it equally natural to talk of the subject's 'knowledge' of where and what the object is, and yet it is at first mysterious to the subject why he is inclined to answer as he does. Confidence comes only with repeated success, or with explanation by someone else of what is to be expected. Yet, even if confidence were strangely present from the first, it would be different in kind from normal or primary perceptual belief. The occurrence or possibility of forms of perception without sensation does not show that sensation is an inessential feature of perception or perceptual knowledge in general, but only that *secondary* perception is possible. It is only because some perceptual knowledge is perspicuously grounded on sensation, some perception is primary, that we have any perceptual knowledge at all.

It may help towards explaining the present argument, and the significance of the distinction between primary and secondary knowledge, to consider a likely objection. The objector in question may admit that self-conscious knowledge plays an important part in the relatively more sophisticated aspects of human life and culture, but claim that at a more animal level we get by on the basis of reliably acquired beliefs without our being reflexively aware of their origins in the exercise of our faculties or in their objects' impinging on us. Since the concept of knowledge extends to such beliefs, it would seem that they must

contain all that is essential to knowledge, and that 'evidence' is icing on the cake. At least, the objector may argue, we would need very strong reason for abandoning this natural conclusion in favour of the notion that 'evident' knowledge is paradigmatic or primary. Yet there is reason to suppose (the objection may continue) that, so far from our possession of 'secondary' knowledge being in any sense parasitic on our possession of 'primary' knowledge, the latter is simply a more sophisticated and, no doubt, evolutionarily late-coming form of knowledge which could not have come about unless preceded by the 'secondary' knowledge which is still good enough for us much of the time. Indeed, it might be said, animals are not self-conscious, but can still properly be said to have knowledge. 'Secondary' knowledge is possible without 'primary' knowledge.

This objection recognizes the role that the distinction between primary and secondary knowledge plays as a methodological alternative to the assumption that there are necessary and sufficient conditions present in each case of knowledge, but misinterprets the conception of 'perspicuousness' or 'evidence' involved in that distinction. The possession of perspicuous knowledge is not as such a sophisticated achievement of human beings, since a distinction can be drawn for animals too, in so far as they can be said to have knowledge, between primary and secondary knowledge. The distinction can be more easily appreciated in the case of human beings, if only because it is itself apparent to self-conscious reflection. Yet is not necessary to be able to reflect self-consciously on one's possession of knowledge in order to have primary knowledge: ordinary perceptual knowledge will do. The real doubt whether animals have memory relates to the question how far they have primary memory, an awareness of the past as past: it is not simply a doubt as to whether they retain knowledge acquired in sense-experience, which they obviously do. The present suggestion is that, if they lack primary memory, their retained beliefs cannot constitute knowledge of the past, whether we call those beliefs 'memories' or (when things have not changed) simply count them as knowledge of the way things are. If we count such beliefs as knowledge of the way things have been, then the past tense should not be supposed to play a role in expressing the content of the belief. Yet, in any case, in order to suppose a being possessing secondary knowledge absolutely without primary knowledge, it would be necessary to suppose a being without ordinary perceptual knowledge. No animal with a serious claim to possess knowledge fits that description. If it is said that a computer suitably hooked up with its environment can possess knowledge without primary knowledge, that argument can be stood on its head. The reason why, despite its rapport with the world, such a computer does not know anything, is that it has no primary knowledge.

Although Locke allowed *only* primary knowledge to count as knowledge, it seems that he did not allow *all* primary knowledge to count as knowledge: 'belief' or 'judgement' is excluded not because it is not consciously grounded, but because, if grounded, it is grounded on 'something extraneous'. It is clear that

what Locke calls 'belief' may possess something at least analogous to the 'perspicuity' or 'evidence' of primary perceptual knowledge or *a priori* understanding. It does so when there is full consciousness of adequate grounds of belief and of their force, so that there is no element of habit or 'hunch' about our believing. It is no doubt because of this analogy that Locke himself occasionally used 'evidence' for the force of 'apparent probability', and that we do so still.¹⁷³ At the same time he needed to hold that such evidence, being fallible, is irremediably second-class and different in kind even from the evidence of sense, not to speak of the 'irresistible light of self-evidence, or...the force of Demonstration'. Locke is wrong because no faculty is infallible. It is therefore unsurprising that the analogy of sight extends naturally to inference, e.g. to the detective's 'seeing' at last who committed the crime. As the paradigm of sense-perception itself makes clear enough, the perspicuity of primary knowledge need not be understood as an infallible criterion of truth. What is important is that it is not therefore reduced to mere conviction.

What, then, is the difference between the 'light of nature' and the 'light from heaven' whose advocates Locke derided as men merely 'sure, because they are sure'? The answer is that the purported 'evidence' of divine inspiration is disjoined from other characteristic features of knowledge, such as the possession of a ground. It should not by now be surprising that what I have called the perspicuity of primary knowledge and its being well-grounded have each to be explained with reference to the other. The ground of a belief, in order to serve as a ground, must be open to the believer, while that openness constitutes the perspicuity of a grounded belief. Like its grounds, the perspicuity of primary knowledge relates essentially to the content of the belief, whereas the 'light from heaven' may attach equally well to any belief, irrespective of its content. As Locke put his case for distinguishing in favour of the light of nature,

Light, true Light in the Mind is, or can be nothing else but the Evidence of the Truth of any Proposition; and if it be not a self-evident Proposition, all the Light it has, or can have, is from the clearness and validity of those Proofs, upon which it is received.¹⁷⁴

The close connection between perspicuity and grounds does not, however, mean that they can be adequately regarded as a single characteristic of knowledge, or that the notion of grounds can do by itself the work necessary in the philosophical explanation of knowledge. Indeed, they can to an extent be split apart. In some cases of secondary knowledge, the belief may be grounded, but not perspicuously so, if the grounds have been forgotten. Or one might remember the proof of a proposition, and that it is the proof, while being no longer able to follow it and 'see' that it is a proof. On the other hand, in the case of Locke's 'intuitive knowledge' it does not seem that our belief has in any natural sense a ground (although it is not groundless), yet it is of course just here that he put most weight on the metaphors of sight and light, and on the notion of

evidence or self-evidence. There is a lack of mystery for most of us about our strong inclination to believe that $7+5=12$ which many present-day philosophers, including Wittgenstein, have felt constrained to treat as if it were the mere consequence of that inclination's being a familiar feature of our natural history. Yet such flat naturalism misses something essential to knowledge which is captured by the traditional metaphors. Most of us believe that, *if A is greater than B and B than C, then A is greater than C* just because we 'see why' it must be so: we understand and grasp the matter well. More needs to be (and will be) said, but there is good enough reason to include what I have called 'perspicuity' among the characteristics of primary knowledge: a feature neither fully separate and distinct from other characteristics, nor reducible to them.

Here it may be objected that the proposal is simply circular, if 'perspicuity' itself needs explaining in terms of an awareness or understanding of why we believe, not to speak of 'knowing how we know'. If we understand what awareness and understanding are, why should we need an explanation of knowledge? In effect this objection is less to a particular item on our list of characteristics of knowledge than to the whole project of philosophical explanation. Take the discussion of the causality of knowledge, in which it was shown that in order to characterize the kind of causality appropriate to knowledge we need the notion of a dependable *cognitive* faculty: it is self-evident that such a notion cannot be possessed independently of an understanding of what cognition is, and so of knowledge itself. No more can it be said that the notions of justification, grounding or truth itself are logically *prior* to the notion of knowledge. In the next section the same will be argued even of the notion of belief. What this seemingly radical objection shows, however, is not that it is worthless to attempt a philosophical explanation of what knowledge is, but that the purpose of such explanation is not to reduce the notion of knowledge to a set of elements quite different in kind from itself. Its purpose is to reveal order, and the place of *knowledge* in an ordered set of notions; and to do so in a way that explains how it is appropriate to have such notions. The present purpose has also been to explain specific philosophical conflict, and systematically to relate the order proposed by others (in particular, of course, by Locke) to the order here proposed.

16

Belief and rationality

Locke's account of belief, even on his own narrow and technical conception of it, is, as we have seen, torn between a variety of theoretical presuppositions and demands. He adhered to a traditional view of affirmation or judgement as the predicative joining of concepts, understanding judgements of probability as the operation of a cognitive faculty which naturally adjusts its output to the grounds of assent. At the same time he was grappling with the problem of error and irrational belief, with the complication that error and irrationality can be both blameworthy and ascribable to the passions. The present section will offer a very brief philosophical assessment of these issues and their inter-relations.

Broadly, two historic models for belief have been considered in the present work. As well as the traditional predicative model adopted by Locke, there is the view adopted by Descartes, that belief is an attitude taken towards a proposition or content which is capable of being the object of some other attitude, such as desire or intention, or of being barely and neutrally held in mind. The latter model can seem vastly preferable but, as it has been seen, it is not without difficulty. First there is the problem of non-verbal belief. A cat looking for its kittens or an infant looking for its toy may intelligibly be assigned some inclination to believe, now that what it seeks is in the cupboard, now that it is under the table (or wherever). Yet we can hardly suppose that the seeker has a proposition in mind with respect to which it can judiciously adopt an attitude or suspend belief, or to which it can assign a probability. The power so to reflect on a propositional content presupposes the power to articulate the proposition in language. Yet an analysis of belief as an attitude towards a propositional content may seem to imply that wherever there is belief there is just such a content capable of neutral entertainment. So it would seem to follow that, after all, the cat and the infant cannot have beliefs, and we have a rather quick proof that beliefs, desires, intentions and the like are impossible without language. The conclusion has had recent advocates, but is a paradox that it would be desirable to avoid. The 'Cartesian' model, if it is to do justice to the continuities between linguistic and pre-linguistic mentality, needs to include some account of how a being without language can hold a content in mind prior to belief. That is not out of the question if all that is essentially involved in belief is some form of mental

representation, not necessarily in the full sense prepositional, together with some sort of acceptance of that representation. If, as it has been suggested above, sensory representation can be regarded as the most primitive form of mental representation, and at any rate provides the paradigm which is about the easiest for us to grasp, then we are also provided with a paradigm of 'acceptance': we commonly accept the deliverances of the senses (we normally believe our eyes, for example) but sometimes we do not. So an animal confused by contrary sensory representations would be faced with a choice at least analogous to the judicious choice a human being might make between rival hypotheses. The 'Cartesian' model then needs only to suppose some other form of pre-linguistic representation in order to allow that an animal or infant can have beliefs about what is not represented in its current sense-experience.

That is not to say that there is no problem about 'acceptance'. The present discussion, however, will be restricted to questions more directly related to Locke's argument and, very probably, to his reasons for rejecting Descartes' account of belief. To what extent can belief be voluntary? How are reason and the passions related as motives of belief? Is anything to be gained by thinking of belief as the exercise of a faculty of judgement?

The interest of Locke's discussion of the first of these questions lies partly in the ambivalence of his answer, which lacks the sweeping and uncompromising (if not altogether unsubtle) character of the analogy which Descartes had drawn between belief and action, error and sin. It may be helpful first to consider some of the objections which can be brought against that analogy, beginning with the crude question as to how on earth it is supposed to be within my power to believe that $7+5=13$, or that this pen in my hand is a snake. Descartes himself in effect considered just this objection in what for him was its strongest possible form, with particular respect to his perception of his own existence. His answer was that, although in some sense he could not but believe that he existed, his belief was nevertheless free in that it was not compelled by an external cause. He believed simply because his assent was motivated by the light of reason: 'I believed it with all the greater freedom and spontaneity, the less I was indifferent towards it.'¹⁷⁵ The thought behind this answer involves a metaphysical model with roots going back to Plato, together with a piece of common sense. According to the former, the soul chooses freely when its choice is determined by its own nature. Since the soul is essentially rational, and the passions and non-rational elements in human nature are due to the soul's unfortunate entanglement with the flesh, the more that a choice is determined by reason, the more it is self-determined and free. The conclusion that less than wholly rational choice is less than wholly free choice is not itself commonsensical, but it is a common-sense or intuitive point that reasons do not as such limit freedom. I have good reasons of many kinds which determine that I am not about to walk around Oxford smashing windows, but their existence does not limit my freedom to do so and take the consequences. Descartes' claim was that, analogously, conclusive

reasons for a belief will determine a rational person's assent, but do not render that assent less free.

That there is something wrong with the analogy is not difficult to see. The notion of will becomes more than usually mysterious when the opposition between success and failure can get so little grip. To explain assent as a form of volition is to say that it is something like choosing or trying, but it can hardly be choosing or trying to believe, since to give mental assent just is to believe. The structure of action, then, is quite different from the structure of belief, and correspondingly, a reason for belief is very different from a reason for action. That appears from the possible contrast between 'grounds' or 'reasons' for a belief and 'motives' for holding it, a distinction which has no analogue in the case of action. If someone offers me a large sum of money to go window-smashing, or even dares me to do so, they have given me a reason like any other which may rationally influence my decision. On the other hand, 'I dare you to believe that you have a snake in your hand' is nonsense. Yet, at the same time, as Locke was so well aware, interests and passions do often determine our beliefs. They are motives as opposed to reasons because of their lack of connection with what is believed. They do not stand to the proposition believed as ground to consequent. Roughly, they are directed towards the prepositional attitude in the circumstances, rather than towards the intentional content of the attitude. But that itself makes them quite unlike motives or reasons for action, for determinants of belief thus dissociated from its prepositional content cannot be acknowledged by us as believers. By that I do not mean that we do not acknowledge them because they are disreputable, but that we cannot in principle acknowledge them even to ourselves. That is why it is senseless to invite me to acknowledge the expectation of financial reward or eternal happiness as the determinant, here and now, of a belief in the virgin birth, or a belief that my pen is a snake. 'I believe that she is dishonest because I want to be revenged' is a paradox which effectively withdraws the initial expression of belief, and is hardly intelligible except as the recognition of the ill-foundedness of a belief held prior to the time of speaking. Someone who really does believe from such a motive will rationalize the belief, since we have to see our own beliefs as well grounded. As Locke put it, 'There is no Body in the Commonwealth of Learning, who does not profess himself a lover of Truth: and there is not a rational Creature that would not take it amiss to be thought otherwise of.'¹⁷⁶ 'Love of Truth', indeed, is one 'passion' which we can admit to as a motive of our present belief just because in the nature of the case it cannot determine belief except through our apprehension of grounds or reasons. Other passions or interests must determine belief indirectly and more or less blindly, through what we call wishful thinking, self-deception and the like. In so far as we doubt that grounds wholly determine our belief, so far is our belief itself subjectively insecure.

Passmore proposes, in the account of Locke's argument discussed in [chapter 13](#) above, that Locke inconsistently wavered between an 'intellectualist' conception of belief as the exercise of a rational faculty of judgement and a view

of it as motivated by passion in the same way as voluntary action. Yet the tension Passmore perceives is endemic to the topic itself, to what Locke was struggling to capture. According to Passmore, Locke eventually settled for an account of the second kind, one of which Passmore himself seems surprisingly inclined to approve. Yet any such account would be incoherent and absurd. In general, we are necessarily saddled with an 'intellectualist' view of our own beliefs, and are committed to their being well-grounded products of our faculties. The love of truth is not a passion like any other but, as Locke clearly indicated, has an exclusive authority in virtue of the authority of 'evidence'. When other passions influence our belief, unreason must pay its tribute to reason in the form of rationalization: or, rather, for the belief to be possible reason must collude with unreason, our cognitive faculties with our emotions and desires. Motivated belief, for all that it owes to emotion or interest, must also be in part a product of our intelligence and powers of interpretation. As such a product it might even, at the extreme, involve a distortion of the deliverances of the senses.

The impossibility of credence without at least seeming credentials explains our intuitive discomfort with the sort of naturalistic scepticism to be found in Hume. We are invited to recognize the ungroundedness (and worse) of our ordinary beliefs, but at the same time to acquiesce in them as the inevitable product of human nature. It is unsurprising that Hume found in himself alternate inclinations to lose hold of the beliefs or to dismiss the sceptical arguments. Yet behind his sceptical recommendation of natural belief lay a view of our relationship to reality which he often hinted at, but on his own principles could not overtly promulgate. That is the view that the operations of the imagination which give rise to our beliefs about the world, although entirely irrational in themselves, nevertheless lead to beliefs which in some unimaginable way correspond systematically to how things really are. Although he insisted, for example, that the idea of causal necessity is an idea irrationally, if naturally, projected by us onto the phenomena, yet at the same time he allowed himself to allude to 'secret connections' and a 'secret nature' of things in virtue which the phenomena take the course they do, and to a 'secret union' of soul and body which underlies the blankly unintelligible operations of the will.¹⁷⁷ In effect Hume was proposing, not that we have no knowledge, but that all our knowledge is a sort of secondary knowledge, lacking perspicuous grounds. Yet, on his principles and on this very model of our relationship to the world, we can have no assurance that sense and imagination do after all afford any kind of contact with a reality of which we never have primary knowledge.

It is as much as anything the sheer difficulty of supplying a model for the collusion of the cognitive faculties with motives, and for the necessary indirectness of the influence of interest and passion on belief, which caused Locke's problems. That difficulty is still with us, although nothing is more familiar than the phenomenon. What seems clear is that no simple two-stage model for judgement will serve the purpose. A standard objection to Descartes' account is that, if we suppose that the understanding presents grounds and

arguments as 'clear and distinct', then there is nothing left for the will to perform, since the selection and assessment of grounds itself constitutes judgement. If there is anything wilful about assent, and if it is through the will that prejudice and passion pervert it, then will, prejudice and passion not only can, but must permeate the whole process of corrupted judgement. A rather similar objection can be brought, as we have seen (and as Locke showed himself at least sometimes uneasily aware), against Locke's rival and somewhat half-hearted attempt to restrict what is wilful in judgement, or subject to ulterior motive, to the stage of directing our gaze and collecting our evidence.

Passmore's distinction between an intellectualist and an emotivist theory of belief is perhaps an exaggeration of the difference between an optimistic and a radically pessimistic view of the collusion between reason and passion. The liberal rationalist will retain, as Locke did, a conception of an objective love of truth which is in principle capable of excluding other passions from the process of judgement altogether. The currently more fashionable pessimist will claim that interest, values and ideology will always and necessarily determine the selection and assessment of evidence. Those who hold the latter view sometimes see a virtue in coupling with their denunciation of the ideal of objectivity an announcement of their own practical aims and values, a confused intellectual performance with just the paradoxical character of a soap manufacturer's declaring that he holds the favourable beliefs about his product that he does because he wants it to sell. The notion of objectivity is not so easily jettisoned. It is a necessary possession of us all as intelligent beings, not just of the enlightened liberal. It is also certain that Locke never wavered in his theoretical adherence to it.

Part III

Perceptual knowledge

Introduction to Part III

Some account of Locke's theory of 'sensitive knowledge', and of its essential connection with his doctrine that simple ideas are natural signs of their external causes, has been given above. Yet the theory deserves closer attention. It comprises, first, a version of the ancient doctrine of the independent authority of the senses and, second, a particular definition of the scope of perceptual knowledge. Both theses flow directly and elegantly from the same neat structure.

In [chapter 18](#) the theory, its point and its relations to other aspects of Locke's thought will be further explored. In [chapter 19](#) an argument will be advanced for a view which is like Locke's in ascribing to sensation a primitive and underived authority as a source of knowledge, but which is without his assumption that a knowledge-delivering faculty is infallible. The proposal is a form of foundationism, but one which takes the foundations to be individually defeasible. It also allows, against Locke, that in sense-perception objects are known to us as more than bare possessors of powers, or than unknown causes of blank sensory effects.

To say that knowledge of the world has foundations in sensory or perceptual knowledge is in effect to say that in sense-perception some knowledge is given prior to both inference and theory. [Chapter 20](#) considers certain arguments or models which are popularly supposed to rule this out, and reasons are given for rejecting them. If perceptual knowledge is foundational, however, it must have an identifiable boundary. In [chapter 21](#) certain criticisms of Locke's model, and of its implications for the boundary of perceptual knowledge, lead into an unorthodox proposal as to where the boundary lies. It is argued that in order to do justice to the scope of perceptual knowledge it is necessary to develop the conception of an integrated sense-field. It is wrong to think of the senses as in general the source of disparate streams of information or content, each discrete from the deliverances of the other senses which it is left to some superior intellectual faculty to relate to one another in constructing knowledge of objects in space. The common objects of different senses are presented as such, and as spatially and causally related to us. In other words, when we both touch and see (or feel 'proprioceptively' as a part of our body, or taste etc.) the same object, the content of perceptual knowledge itself commonly and normally includes its

being the same. There is inter-sensory perceptual knowledge, and there are inter-sensory illusions.

These theoretical proposals involve a notion of intrinsic intentional content, and indeed of consciousness, at odds with much present-day philosophy of mind, although consonant with ordinary ways of thinking about perception. In [chapter 22](#) two current explanations of content, functionalism and an explanation deriving from Wittgenstein, are examined and criticized. Their apparent shortcomings are taken to point to the kind of conception of content adopted in the present argument.

Finally, in [chapter 23](#), Locke's conception of sensations as signs, and some aspects of his treatment of secondary qualities, will be compared with some modern suggestions about the relationship between sensation and the meaning of such terms as 'blue' and 'hot'. Another view of this relationship, consonant with the conception of the content of sensation proposed in the earlier chapters, will be drawn from the comparison.

The authority and limits of ‘sensitive knowledge’

In the ‘actual receiving of ideas from without’ we are, according to Locke, immediately aware, not only of the sensation or idea, but also that ‘something doth exist at that time without us, which causes that Idea in us’ through the senses. Sensitive knowledge is immediate awareness of an idea as a certain sort of effect. Now according to Locke’s general theory of representation or signification, as we have seen, a simple idea is automatically fitted to represent in our thought precisely that ‘something which causes that idea in us’: the idea of white represents the quality white conceived of as an external power. Hence sensitive knowledge is knowledge of the existence of such qualities, whatever actual attributes of things give rise to them. As Locke put it:

whilst I write this, I have, by the Paper affecting my Eyes, that *Idea* produced in my Mind, which whatever Object causes, I call *White*; by which I know, that that Quality or Accident (i.e. whose appearance before my Eyes, always causes that *Idea*) doth really exist, and hath a Being without me. And of this, the greatest assurance I can possibly have, and to which my Faculties can attain, is the Testimony of my Eyes, which are the proper and sole Judges of this thing.¹

The last sentence quoted is a careful and emphatic declaration of the independent authority of the senses: independent, that is to say, of any support that reason might give. The whole passage is in terms strongly reminiscent of ancient empiricism.

The significance of this doctrine in Locke’s time may be better revealed by some comparison with Cartesian theory. For Descartes, as it has been explained above, an idea of sense is ‘material’ for judgement, and capable of being properly or improperly formed. An idea of blue which is formed by the mind’s referring the characteristic sensation to a cause supposed just like it is ‘materially false’, and the corresponding judgement is hasty and erroneous. An idea of blue, on the other hand, formed by the mind’s referring the sensation simply to an unspecified cause, i.e. one specified only as the normal cause of such a sensory effect, is materially true, but contains something which is obscure. Such an idea

was perhaps seen by Descartes, and was certainly seen by his follower Arnauld, as the truly natural and immediate deliverance of the senses. The corresponding judgement is limited in its content and fallible in particular cases, but it is ‘what nature teaches’, the effect of an innate propensity to believe, and is in general useful and reliable. Yet in all of us this first, natural belief has been overlaid by assent to the materially false idea of blue. As the Port Royal *Logic* put it, ‘the mind has added false judgements to what nature reveals to us’. As we should expect, a God-given tendency to believe, although fallible, is not intrinsically deceptive. But there is a third sort of idea and judgement, the only one that is rationally based and so fully armed against scepticism. *Knowledge* of the existence of external objects, as opposed to mere natural belief, comes only after the philosophical reflection, issuing in a clear and distinct idea, which leads us to judge that the sensation is caused, via a mechanical chain of movements stretching from the object to the brain, by the shapes and motions of particles of Cartesian matter. It is indeed only this scientific view of sense-perception which can reassure us, once infected by scepticism, that the second, natural sort of judgement is relatively dependable, and can enable us to understand when we should, and when we should not rely on it.²

Locke’s account of perceptual knowledge can be read as a deliberate and often subtle attempt to be systematically different from Descartes, very broadly in the footsteps of Hobbes and Gassendi. First, he evidently rejected both the scepticism and the attempt to meet it by demonstrative, scientific argument: the former is psychologically impossible and methodologically unnecessary, while the latter dogmatically assumes a questionable speculative hypothesis about mind, matter and the mechanism of sense perception. It is tempting at this point to conclude that Locke was simply telling his readers to be satisfied with Descartes’ ‘teaching of nature’, and that he was commending the candle-light of natural (or, at best, probable) *belief* because the broad sunshine of certain *knowledge* is impossible. In the discussion of the doctrine of degrees of knowledge in [chapter 11](#), above, passages were considered which can seem to support that view of Locke’s intentions; but inspection of the whole argument revealed a different meaning, one which does not in the same way run counter to the distinction between knowledge and belief. Locke was being serious when he said that the senses give us ‘an assurance that *deserves the name of Knowledge*’.³

It is easy to receive the impression that Locke was contradicting himself, sometimes in a single sentence. The senses provide ‘an Evidence, that puts us past doubting’.⁴ Yet our faculties are ‘suited not to the full extent of Being, nor to a perfect, clear, comprehensive Knowledge of things free from all doubt and scruple; but to the preservation of us, in whom they are’, a sentiment which seems to echo the Cartesian view of ‘what nature teaches’ and natural belief. And yet, again, that judgement is immediately glossed with the assertion that the senses ‘serve our purpose well enough, if they will but give us *certain* notice of those Things, which are convenient or inconvenient to us’.⁵ The apparent self-contradiction can, however, be explained. The concession is that the senses do

not give us a scientific understanding of things, for they provide knowledge of existence but not knowledge of essence. The knowledge of existence they give us is open to a certain 'doubt and scruple', but here Locke seems only to have had in mind the general sweeping doubt of philosophical scepticism, a metaphysical scruple which he believed can after examination be rejected absolutely, with derision. His claim was, indeed, that it can be so dismissed without recourse to theoretical knowledge and reasoning: 'For it takes not from the certainty of our Senses, and the *Ideas* we receive by them, that we know not the manner wherein they are produced.'⁶ Thus while it is true and important that what Locke proposed as 'sensitive knowledge' corresponds to Cartesian natural belief, it is no less important that he carefully and deliberately transformed its epistemological status.

In order to understand how Locke arrived at this position it is necessary to bear in mind what has been said before in explanation of the dependable relation of simple ideas to reality. Simple ideas represent or signify their causes. It is that invariable natural relation, and nothing else, which constitutes the essential relation between an idea and its object. For Descartes, as it seems, the representative content of a sensory idea is determined by how the mind relates a datum to a supposed cause, an act which can be done well or ill, perversely, or naturally, or rationally. But for Locke, once the mind uses the datum as a sign (and that seems to be assumed as inevitable), what the idea signifies is itself in effect a datum too. A simple natural sign can only signify its normal cause, whatever that may be.

A nice illustration of the difference between the two theories hinges on a difference on a point of religion. The Port Royal *Logic* contains the Cartesian claim that the certainty of the senses 'arises not only from the senses but from the reasoning which distinguishes when we ought and when we ought not to believe them'. Its authors could go on to argue that the doctrine of the Eucharist does not contradict the senses, since even ordinarily the senses must be interpreted by reasons. In the Eucharist we simply have extraordinary reasons for taking the body of Christ to be behind the appearance of a wafer.⁷ But Locke had another view of the relation of such 'reasons' to the deliverances of sense, so that the independent authority of sense could be brought in the Protestant cause. The Romanist is indoctrinated from childhood with the principle that Church and Pope are infallible:

How is he prepared easily to swallow, not only against all Probability, but even the clear Evidence of his Senses, the Doctrine of *Transubstantiation*? This Principle has such an influence on his Mind that he will believe that to be Flesh, which he sees to be Bread.⁸

It is therefore a potentially misleading feature of the chapter on sensitive knowledge that Locke soon moved into a rebuttal of scepticism which can look rather like a set-piece argument for the existence of external objects, reasoning

designed to transform natural belief into knowledge. This argument appeals to the phenomenal peculiarities of actual perception by contrast with memory and imagination, and to the need for a certain sort of causal process involving a sense-organ or receptor in order to produce anything like the former. It appeals too to our inability to conjure up those all-important effects of the objective world, physical pain and pleasure. A further and familiar consideration is the systematic character of sense-experience, so that what is learned by one sense can be checked by others. There is also said to be an evident difference between our imagining something and our actually bringing it into existence by a physical action: each is in its way dependent on our thought and will, but the product of the latter has thenceforward the intractability of the objective world. There are one or two arguments more, and the whole passage can encourage us to suppose that for Locke, as for many other philosophers, our perceptual beliefs are the conclusions of reasonable inferences based on the character and pattern of subjective experience, and that he was here concerned with the criteria we use in arriving at them.⁹

Locke made it quite clear what he was really doing, however, since the explicit purpose of these arguments was precisely not to supply a rational foundation for perceptual belief, which he took to be already securely founded on the appropriate cognitive faculties, the senses themselves. Still less was it to list the criteria we ordinarily use in perceptual judgements. It was merely to supply what he called 'concurrent reasons' to confirm, in the face of doubt, 'the assurance we have from our Senses themselves'. There is no reason to suppose that Locke had everyday doubt in mind: it is the sceptic who needs special treatment because he goes to the extreme of pretending to doubt whether his natural faculties really are cognitive faculties, delivering knowledge. Thus Locke's first response to scepticism does not even look like an argument for an indirect, speculative interpretation of our sensations *ab initio*, but is something rather more interesting. Coupled with an appeal to the 'evidence' of perceptual knowledge and the impossibility of genuine doubt at a practical level, the point is made that metaphysical mistrust of the senses is mistrust of a basic cognitive faculty, a mistrust which makes the whole concept of knowledge meaningless: 'For we cannot act any thing, but by our Faculties: nor talk of Knowledge it self, but by the help of those Faculties, which are fitted to apprehend even what Knowledge is.' Philosophically educated readers of the *Essay* could hardly have failed to recognize an Epicurean response to scepticism, familiar from Lucretius: 'if [the sceptic] has never before seen anything true in the world, from where does he get his knowledge of what knowing and not knowing are?' For Lucretius, 'the preconception of true has its origin in the senses'. Since reason is the product of the senses, 'if the senses are not true all reason becomes false as well'.¹⁰ Since Locke admitted degrees of evidence superior to that of sensitive knowledge, he could not go so far so fast, yet on his account too reason has its dependence on sense. Our faculties, he implied, stand or fall together. In this he was quite unlike Descartes, for whom the senses uncorrected by reason are intrinsically liable to

illusion, and reason is the providential faculty by which truth may be securely distinguished from falsehood, and which stands in judgement over the senses.

Sensitive knowledge is knowledge of existence, and it is easy to see how, on Locke's causal account of representation, the knowledge that this particular sensation or idea of white is caused by something outside us comes to the same thing as the knowledge that what this idea represents (the quality whiteness) exists. It is less easy to see how such knowledge conforms to the general definition of knowledge, i.e. how it consists in perception that the idea of existence, or 'real existence', 'agrees with' the idea of white.¹¹ Presumably both ideas must be before the mind if we are to perceive that they are related, and yet there seems to be nothing in sense-experience to count as the idea of existence over and above the idea of white. It is not at all surprising, in the context of this difficulty for Locke, that Hume should have come roundly to declare on behalf of his more sceptical theory that there is no idea of the existence of an object distinct from the idea of the object; or that he should correspondingly have characterized both sense impressions and natural perceptual belief in terms of the only phenomenal peculiarity that he was prepared to allow them, the 'force' and 'vivacity' of actual sensation. Nevertheless the point constitutes more of a difficulty for Locke's definition of knowledge and consequential subject-predicate analysis of existential propositions than for his realist account of sense-experience as the immediate, non-inferential awareness of our environment's acting upon us.

More will be said about the senses' authority or autonomy. What about the scope of sensitive knowledge, on Locke's account of it? His model restricts the intrinsic intentionality of sensation to the causality of the idea, the otherwise blank effect: what is given in sensation is that this idea is caused by something 'without us'. Sensitive knowledge is therefore knowledge of sensible qualities just in so far as they are powers to act on us. The expression 'without us' might, however, be taken to imply that we have sensitive knowledge of things' spatial location as well as of their powers to affect us, and such an interpretation might seem to receive support from the closely following remark that in sensation 'I know, that that Quality or Accident (i.e. whose appearance before my Eyes, always causes that *Idea*) doth really exist, and hath a Being without me.' On that interpretation, the spatial position of things 'before my eyes' would come under the scope of sensitive knowledge as more than a bare power to affect us: it would somehow enter into the intrinsic intentional content of sensation on the same level as the ideas' external causality. Now the question whether 'without us' is here to be understood spatially or only, as we might say, meta-physically was doubtless not a question Locke ever directly asked himself; and that it arises at all indicates the tension between his different notions of an idea. Yet it does seem clear that the notion of ideas as signs to which he was appealing throughout his account of sensitive knowledge takes them as what I have called 'blank effects' on the mind. The same notion is implicit in the principle that all simple ideas are true. To suppose that for Locke spatial location and attributes are given in sense-

perception as features of reality will lead us in particular to misrepresent the status of his distinction between primary and secondary qualities, which he regarded as a hypothesis. To draw it, he thought, was to make an 'Excursion into Natural Philosophy',¹² if a pre-eminently reasonable one. In themselves, without that excursion, the senses no more tell us what lies behind the spatial appearances of things, than they tell us what lies behind their colours and smells. Even if ideas of primary qualities do resemble their causes, there is nothing about them which, without argument, could assure us that that is so, and that they are more than blank effects giving knowledge of bare powers.

Another, rather more complicated question about the scope of Locke's 'sensitive knowledge' is this: does knowledge of the existence of sensible qualities constitute knowledge of the existence of the substantial things which possess those qualities? What kind of step is it from 'White exists without me' to 'A white thing exists without me'? Since, on the model for perception which Locke was elaborating, sensible qualities just are the various powers of unitary things or substances to act on the different senses, one would expect the step to have been regarded as minimal, intuitive if not tautological. That is, indeed, how he presented it to Stillingfleet: 'all simple ideas, all sensible qualities, carry with them a supposition...of a substance wherein they inhere'.¹³ To his correspondent Samuel Bold he wrote

I agree with you, that the Ideas of Modes and Actions of Substances are usually in our Minds before the Idea of Substance itself; but in this I differ from you, that I do not think the Ideas of the Operations of Things are antecedent to the Ideas of their Existence; for they must exist before they can in any way affect us to make us sensible of their Operations, and we must suppose them to be before they operate.¹⁴

Nevertheless the passages in the *Essay* which provoked Stillingfleet's objections can be taken to suggest that the 'supposition' is not so immediate and unquestionable as that. In perhaps the most famous passage, Locke had written that, 'not imagining how these simple Ideas can subsist by themselves, we accustom ourselves, to suppose some *Substratum*, wherein they do subsist, and from which they do result, which therefore we call *Substance*'.¹⁵

The apparent conflict can be resolved, if we take it that the supposition which calls upon custom is rather different from the tautology that what acts on us exists. That the latter is in effect encompassed in sensitive knowledge appears from the senses' supplying knowledge not only of particular existence, but of particular coexistence: 'we cannot so far distrust their Testimony, as to doubt, that such Collections of simple Ideas, as we have observed by our Senses to be united together, do really exist together.'¹⁶ Since for qualities to 'exist together' just is for them to be qualities of the same thing, to hold that we possess sensitive knowledge of the coexistence of qualities but lack such knowledge of the existence of substantial things would be self-contradictory. Not that there are no

problems here, but Locke does not seem guilty of that particular incoherence. For the supposition which seems to go beyond sensitive knowledge, the supposition which is involved in the formation of 'our complex ideas of substances' (and of that general idea of a substrate or substance which is an ingredient of such ideas), is a supposition about *universal* coexistences, not *particular* ones. The account of the genesis of ideas of substances (and the idea of substance in general) *begins* with the statement that the mind notices in sensation and reflection that a certain number of simple ideas 'go constantly together'. It is not, it seems, the perception of the existence of a single quality which gives rise to the explicit idea of substance or *substratum*, nor even the perception of coexistence on a particular occasion, but the *repeated* perception of the coexistence of the same qualities: e.g. the yellow colour, weight, malleability, and the rest commonly included in the idea of gold. It is then that we form the idea of something, a unitary unknown, which is responsible for this recurring union of attributes. Just as we can perceive particular identities and differences without the aid of an explicit idea of identity or diversity, so, Locke seems to have thought, we can perceive particular coexistences of qualities (and, *a fortiori*, particular qualities on their own) before forming the explicit idea of a 'substance' to which the qualities belong. That may even be why, in discussing sensitive knowledge of coexistence, Locke referred to a particular man simply as 'a collection of simple ideas'. Only in the formation of universal specific ideas ('*Man, Horse, Sun, Water, Iron*')¹⁷ does a general idea of substance become an explicit sign in the language of thought. And given that, for Locke, in forming such ideas we presuppose that the attributes which we have observed together on particular occasions are universally and dependably conjoined, then it is unsurprising if he took the process to lean on 'custom' and to be attended with a certain risk. All of this may be made clearer in the discussion of substance in Volume II: for the present it may be assumed that Locke saw no significant gap between sensitive knowledge of the existence of a quality on some occasion, and knowledge of the existence of the substantial thing which is then acting on us through the senses. Knowledge of its essence is of course another matter.

There is here an important analogy between Locke's explanation of the formation of the explicit idea of a substance or substrate, and his explanation of the genesis of the related ideas of power and of cause and effect. For his theory of sense perception and his understanding of the notion of power seem to interpenetrate each other. On the one hand, the observed regular changes which give rise to the explicit idea of power include those which the mind takes notice of, 'reflecting... on what passes within it self, and observing a constant change of its *Ideas*, sometimes by the impression of outward Objects on the Senses'. Thus the acquisition of the idea of power is not entirely distinct from the apprehension of an objective world, a world represented by our ideas. On the other hand, a notion of power is somehow present in the relation of representation itself, for simple ideas represent powers. Locke's position seems to have been that, although in some sense involved in the representative relation,

the notion of power is, as it were, absorbed by it. In other words, our idea of a sensible quality does not *include* the idea of power: the idea of power is not an explicit component of the existential proposition constituting sensitive knowledge that blue now exists without us. The idea of power becomes explicit only with the second-order recognition that one thing is bringing about a change in another perceived thing, a change which is itself the object of first-order sensitive knowledge:

The Mind...concluding from what it has so constantly observed to have been, that the like Changes will for the future be made, in the same things, by like Agents, and by the like ways, considers in one thing the possibility of having any of its simple *Ideas* changed, and in another the possibility of making that change; and so comes by that *Idea* which we call *Power*. Thus we say,...the Sun has a *power* to blanch Wax, and Wax a *power* to be blanched by the Sun, whereby the Yellowness is destroy'd, and Whiteness made to exist in its room. In which, and the like Cases, the *Power* we consider is in reference to the change of perceivable Ideas.¹⁸

In such perceptually second-order ascriptions of power, the idea of power has reference to what we do not observe, to whatever in the agent and patient lies behind the observed change in the latter. In itself it has no positive content, since the positive content of the idea of a particular power is supplied by the particular observable effect. Thus for Locke the idea of power operates in everyday thought as a sort of dummy concept or place-marker, a product of our ignorance of the intrinsic attributes of things. By its means the relation of signification is, so to speak, made capable of a certain extension. The idea received from the observed whitening of wax comes, by being added to the idea of power, to signify or represent not only the change or novel attribute of the wax, its coming to be white, but also those underlying and pre-existent intrinsic or structural attributes of both the sun and wax in virtue of which wax regularly turns white whenever it is appropriately exposed to the sun. What is relevant to the present discussion is that this second-order representation rests on something going beyond particular perceptions, namely the perception of constant and dependable change in recurrent observable circumstances. It is therefore in contrast to the immediate awareness of things' acting on us in sensitive knowledge. Indeed no such regular connections as that between exposure to the sun and the whitening of wax, or the application of heat and the melting of gold, could possibly be identified prior to the perception of an objective world: that is to say, of sun, wax, heat and gold.

It can seem surprising, especially in the light of later scepticism but also against the background of the doctrine of innate ideas, that Locke's treatment of the ideas of power and of cause and effect are separate, that they come so late in his argument and that they are so unemphatic. Most of the chapter on power concerns the theory of human action and freedom, while the brief discussion of

cause and effect is tucked away in what is hardly more than an appendix to the chapter on ideas of relations. Both discussions, taken as accounts of the genesis of an idea from experience, can seem unsatisfying and to invite the charge of circularity. For the suggestion that we cannot reasonably avoid taking discontinuous but 'regular' observable behaviour as an indication of underlying, enduring, causally operative structure, while it may be acceptable, seems to presuppose not only the concept of causality, but that there is a causal order. It assumes that events do not just happen, and that experience is experience of, and within, a law-governed universe. Yet Locke did not see the idea of causality as prior to the idea of power, for the former is explained in just the same terms as the latter:

In the notice, that our Senses take of the constant Vicissitude of Things, we cannot but observe, that several particular, both Qualities, and Substances, begin to exist; and that they receive this their Existence, from the due Application and Operation of some other Being. From this Observation, we get our Ideas of *Cause* and *Effect*. . . . Thus finding, that in that Substance which we call Wax, Fluidity, which is a simple *Idea*, that was not in it before, is constantly produced by the Application of a certain degree of Heat, we call the simple Idea of Heat, in relation to Fluidity in Wax, the *Cause* of it, and Fluidity the *Effect*.¹⁹

Once again observed regularity is what prompts the idea, when experience of 'constant vicissitude' compels us to accept certain hypotheticals as reliable. Yet the idea has reference beyond what is observable, for the cause normally produces its effect 'working by insensible ways'. Locke called the otherwise unknown but rationally presumed mechanism linking cause and effect the '*modus operandi*' or 'manner of operation'. But simply 'to have the *Idea* of *Cause* and *Effect*, it suffices to consider any simple *Idea*, or Substance, as beginning to exist, by the Operation of some other, without knowing the manner of that Operation'.²⁰ Thus the idea of cause and effect is explained in the same way as the idea of power, namely as a means by which a relationship in itself unknown is nevertheless represented in thought. The same goes for specific causal notions, the ideas of 'actions' such as *freezing*: for '*many words which seem to express some Action, signify nothing of the Action, or Modus Operandi* at all, *but* barely the *effect*, with some circumstances of the Subject wrought on, or Cause operating'.²¹ In all these cases the argument seems open to the charge of circularity: how does experience assure us that there are absolutely law-governed processes and mechanisms underlying the only relatively regular sequences in the world as we experience it? The critic may feel that the real problem is not, as Locke seems to have thought, the problem of how such mechanisms are represented in thought, but the problem of how we know or arrive at the conception that they are there to be represented. Even if it is accepted as a necessary truth that everything has a cause, it can hardly be

supposed that the mind should appeal to the principle of universal causation before it has acquired the idea of causation. Moreover, there may seem to be a discrepancy between the role ascribed to the observation of constant conjunctions in the generation of the causal ideas and the doctrine of sensitive knowledge as the immediate awareness of a causal relation on a single occasion.

Such criticisms can perhaps be put into perspective if we draw the right conclusions from the evidently limited purpose and explicit assumptions of Locke's account of the formation of the ideas of power and cause. For it is unreasonable to suppose that he was there intending to offer an explanation from scratch of how we come, in the most general sense, to think in causal terms. In effect everything about his 'historical' epistemology²² and theory of representation seems designed to convince us that we could not take the first steps in our knowledge of the world, nor even, perhaps, could we in any real sense think, without some awareness of causal relations between ourselves and other things. In that, whatever criticisms might be brought against his theory of representation, he was no doubt right. It seems then that we should read his accounts of the causal ideas as what even on the face of it they appear to be: i.e. as explanations of how we come to employ certain dummies or variables in our thought which enable us in very specific ways to think and have limited knowledge of what in itself falls outside the scope of experience. These place-markers arise from a response of the mind to experience which is as natural and rational as the use of simple ideas as signs of their causes: 'Since whatever Change is observed, the Mind must collect a Power somewhere, able to make that Change, as well as a possibility in the thing it self to receive it.'²³ From Locke's point of view that response need not be supposed to depend on an explicitly or (which made no sense to him) implicitly formulated thought that nature is law-governed. For, as we have seen, what Cartesians and others would explain in terms of innate principles or concepts, Locke characteristically explained as the simple exercise of our mental faculties. Accordingly, the use of simple ideas as natural signs of their regular causes was not for him a result of any reasoning such as that which employs the principle of universal causality, or appeals to the rule that like effects in like circumstances have like causes. It is just thinking and experience at the simplest and most basic level.

Does perceptual knowledge have independent authority?

The view that the senses give us knowledge of the things which surround us is a natural one, but the view that they never deceive us is not. Yet Locke's theoretical justification of the former implies the latter. What is fundamentally at fault is no doubt his division of our cognitive faculties into those which deliver authoritative knowledge, and those which issue in fallible judgement or opinion: '*Error* is not a Fault of our Knowledge, but a mistake of our Judgement giving Assent to that, which is not true.'²⁴ From broadly similar assumptions Descartes had argued that, since the senses may lead us astray, their deliverances require the validation of reason before they can be supposed to give rise to knowledge. Locke, in maintaining their independent authority, evidently felt obliged to discount the possibility of their leading us astray.

It is a remarkable, perhaps astonishing, fact that nothing comes nearer in the *Essay* to being a discussion of sensory illusion than the famous account of 'how the same Water, at the same time, may produce the *Idea* of Cold by one Hand, and of Heat by the other: Whereas it is impossible, that the same Water, if those *Ideas* were really in it, should at the same time be both Hot and Cold'. Locke's explanation of this phenomenon is strongly reminiscent of Gassendi's in his argument that impressions are always true. It is that ideas of heat or cold correspond, not to the absolute degree of motion of the particles either in objects or in our organs of touch, but to 'the increase or diminution' of motion in the latter 'caused by the Corpuscles of any other Body'.²⁵ Since a simple idea represents (is a sign of) its normal cause, Locke was not so much explaining a contradiction between different deliverances of the senses as explaining it away: he represented both sensations as in effect veridical. The existence of this Epicurean line of thought may help to explain the neglect of illusion in the discussion of sensitive knowledge, but another factor, perhaps, was Descartes' having moved the centre of the dispute with the sceptic from the ancient topics of sticks that look bent and round towers that look square to the possibility of a coherent and interminable dream or deception in the presence of nothing at all.

Yet Descartes did say something about illusions in his debate with the sceptic, and it may still seem surprising that Locke did not employ any of the other available responses to sceptical arguments from illusion. Gassendi, for example,

in his reconstruction of Epicurus, had employed the argument that illusions cannot undermine the senses' authority because it is only by reference to the suffrage of the senses as a whole that some experiences are judged misleading or untrustworthy.²⁶ In other words, we appeal to a range of experience in order to judge that a particular appearance is a sign of something in the conditions of perception rather than something in the object. Given the echoes of ancient empiricism in the *Essay* it seems incredible that Locke was unaware of such arguments, but it is possible that he was dissatisfied with them. Perhaps he objected to a distinction too much like Descartes' own between perception and consequent judgement. Moreover, although the common agreement or suffrage of the senses provides a criterion for such judgements which, unlike Descartes', derives from the senses themselves, Locke seems to have felt that each *particular* sense-experience has its own intrinsic evidence and authority, and that to fall back on such a general criterion would be to concede too much to the sceptic. It is therefore, perhaps, not so surprising that he concentrated on the silliness of those who go so far as to deny all significance to the experienced difference between actual sensation and the 'bare imagination' of dreams and day-dreams. Insistence that there could in principle be dream-experience subjectively just like actual experience is, on Locke's view, no different in kind from the metaphysical and unreal insistence that there could be contradictions indistinguishable from evident necessity. Either claim would merely undermine trust in our 'apprehensive faculties' in general and so, if effective, would make argument with the sceptic, and his argument with us, pointless and ludicrous.

Any modern reader of Locke who feels the intuitive force of his claim that the senses deliver immediate, uninferred knowledge in their own right is faced with the task of finding a justification for it quite different from the one to be found in the *Essay*. Locke's neat causal theory of representation, his notion of sensory ideas as blank effects and his conception of them as the constituent terms of a natural language of thought are hardly available to us today as materials for a structured account of perceptual knowledge capable of explaining the 'assurance we have from our senses themselves'. It is clear that a part of the task is to peel off a notion of independent *authority* from the notion of *infallibility*. What follows is an attempt to show how such a thing is possible, employing a number of conclusions which have been arrived at above. It is controversial at every point, and some rival approaches will be considered briefly at a later stage.

One reason for distinguishing, as Locke did not, between perception and perceptual knowledge is that it is possible to see that something is so without believing our eyes. Admittedly, our saying without qualification that James saw that there was an elephant on the lawn carries the implication that James believed his eyes, but the implication is conversational, not logical: if he had not believed his eyes, we would have said so. The notion of perception, broadly speaking, is the notion of successful sensation. Roughly, perceptual knowledge is related to perception as perceptual belief is related to sensation. Perceptual belief is belief which is grounded on sensation, a relation which is at the same time causal and

intentional. That is to say, the belief derives its content from the content of the sensation which causes it: perceptual beliefs are those beliefs we have just because and in so far as we take our sensory states to be veridical. This derivation is not like the derivation of conclusion from premise: rather it consists in our simply and naturally taking what sensation gives. Hence we take what we take immediately and without inference.

‘Successful’ sensation is not just veridical sensation. For perception to take place, a sensory state must be an effect of the state of affairs which satisfies it, and the causal relation must be of a certain kind. Drawing on the account of knowledge given in [chapter 15](#), above, we can say that the sensory state must be veridical in virtue of the functioning of an appropriate faculty or sense-mechanism. Perception can be regarded as the process issuing in sensory states which are in just that way ‘non-accidentally’ veridical: i.e. as the successful functioning of the senses. This notion of perception does not exactly conform to our ordinary ways of talking about ‘seeing’, ‘hearing’, ‘perceiving’ and the like even as those terms are employed for sense-perception. For one thing, as Descartes remarked, we often say that we see what we really infer from what we see. But there is another more radical point which relates to the reference of sensation. As it has been explained above, it is not necessary that my perceptual state be veridical in order for it to be the case that, in the ordinary sense, I see or perceive my shirt. For in perceiving it, I might misperceive it as a face. But, as it was concluded, even misperception requires a measure of sensory success: the face I seem to see must be visually correspondent with the shirt I see. I see the shirt as a face.²⁷ For the present purpose, however, this qualification is not very important, although it is one to which the argument will return.

Nevertheless, the question of the senses’ authority can be approached through the topic of perceptual error, taken to be false (or accidentally true) perceptual belief grounded on falsidical sensation, or illusion. Perceptual error is often avoided at the time if we have reason to believe that things are not to be taken as they have been presented to the senses. Such reason need not, as a traditional approach to illusion has assumed, include reasons for supposing that our senses are malfunctioning or that circumstances are abnormal. The standard optical illusions, for example, arise precisely when our sight is operating normally in normal conditions, given certain types of object. But however that may be, the familiar power of ‘reason’ over natural perceptual belief has led many philosophers to follow Descartes in requiring not only that natural belief should sometimes be rejected in the face of reason, but that strictly it should only be accepted when endorsed by reason. A version of this assumption is the popular view that our beliefs about the external world, including the belief that any physical objects exist at all, form a vast quasi-scientific explanatory hypothesis about our sensations which our sensations on the whole confirm. Epistemologically respectable trust in the senses on any particular occasion would then be extrinsically grounded, derivative from trust in that hypothesis. Yet any advocate of the thesis that the *particular* deliverances of the senses

possess independent authority must hold, not only that perceptual belief is not normally supported by reasoning (which might have been just because we are creatures of habit), but that it is normally inappropriate to require that it should be so supported when it is not. The exceptions arise only when there are countervailing considerations. The deliverances of the senses are defeasible, but have *prima facie* authority. We do not require that an eye-witness bring evidence that his eyes were not deceiving him, unless there is some reason to suppose that they were.

If that authority, ‘the clear evidence of the senses’, is supported by ‘concurrent reasons’, all well and good. We may be more certain that something is as we perceive it to be, if it appears to be as we should otherwise expect it to be. Another matter of degree will be the quality and extent of the individual perceptual experience: a long and careful observation, other things being equal, will afford better grounding than something less. A doctrine of the independent authority of the senses may deny neither of these things. What it must assert is, firstly, that the natural and normal acceptance of the deliverances of sense is not in fact based on speculative reasoning or inference or theory; and secondly and more importantly, that such acceptance is nevertheless epistemologically respectable, at least *prima facie*. In other words a doctrine of authority must refute the Cartesian view that, unless reasons for it can be given, acceptance of what the senses present is merely natural and normal belief, and not knowledge. Now here it is relevant to bear in mind Locke’s appeal to the thought that our senses are after all basic cognitive faculties on a par with intuition, rather than the mere suppliers of mute grist for the intellect’s mill. The natural, even if not (as Locke seemed to think) inevitable issue of the processes of sense perception is knowledge of the environment and of the subject’s place in it; and that, it is arguable, should be enough. Were there no process, independent of reasoning, the natural outcome of which is knowledge of the physical world, we could *have* no knowledge of the physical world. If the world were not *presented* to us by a natural process at least broadly like that of our actual and imperfect animal sensation, then ‘reason’ could never have invented or postulated or ‘posited’ such a world, whatever mute data it can be supposed (*per impossibile*) to have received. Indeed, the notion of reason operating without sense is an absurdity: the first steps in knowledge are necessarily primitive instances of perceptual knowledge.

This position can perhaps be supported by the consideration of a species of ‘unreasonable’ perceptual belief, i.e. the case of someone who continues to trust his senses in spite of a very strong reason for doubt, and who is not deceived. Such a person did perceive what he believes he perceived, his sensation was non-accidentally veridical. Descartes defined error as rash and unreasonable belief (which can nevertheless be true), but, however strong the reasons against our stubborn observer’s judgement, he is certainly not guilty of *perceptual* error. But does he have perceptual knowledge? On my account of perceptual knowledge it

seems that he would have it. That is not an unreasonable claim and, with some minor qualifications, it seems to be an important truth.

The most natural illustrations of that truth involve only circumstantial evidence of misperception, although, as such, that evidence can be indefinitely strong. First, let us imagine that a military observer on the lookout for enemy aeroplanes has had the visual impression of a plane entering a small cloud, but that this experience was unaccompanied by the usual blips on the radar screen, noise or supporting observations of others who were looking in the same direction. Moreover nothing comes out of the cloud. The 'rational' judgement is that his eyes deceived him. But, second, let us suppose that he did in fact see a plane, a silent secret weapon still hovering undetectably in the cloud. If he came to share the rational incredulity of his fellows he would lack (or at least cease to have) perceptual knowledge just because he would lack belief. But equally, contrary to his new belief, he would not have been the victim of perceptual error, since his senses have not, in fact, deceived him. Finally, however, let us make the different supposition that, after all, in the face of all the reasons he refuses to be persuaded and continues to believe that he saw what he in fact did see. Now if it would have been perversely stubborn and irrational of him to have held out had he been wrong, then we would need to suppose that he is perversely stubborn and irrational in holding out when he is right, at least if it is conceded that illusion can be subjectively indistinguishable from perception. Nevertheless it seems not improper to attribute knowledge, perceptual knowledge, to such a person, i.e. to condone his plumping for 'sense' in preference to 'reason'. He knew that there was something there because he saw it. If that is correct, the principle that sensitive knowledge has independent authority receives some endorsement. For my example would suggest that, when it comes to the question whether someone's belief is well-grounded, or on the contrary perverse and unreasonable (i.e. 'error' in Descartes' sense), then their response to the deliverances of their senses is not just one item to be weighed in the scale together with their response to all the other considerations which exist for them on the one side or the other. But that is what it would be, if the deliverances of the senses had in general to be the subject of reasonable interpretation. The implication is that the senses supply the fundamental grounding of our knowledge of the world: what they present as the case, the intentional content of sensation, is, in itself and without reasoning, evidence of what is the case. In other words, in sense perception facts are *evident* to us.

The argument here becomes continuous with the discussion of primary and secondary knowledge entered upon in [chapter 15](#) above. It is the dependence of the content of the belief on the content of sensation which makes perceptual belief more than a natural hunch or inclination. The consciousness which is peculiar and necessary for experiential knowledge and belief involves the consciousness of this derivation and so, as we shall see, of the derivation of the belief from the state of affairs perceived. There is no mystery to the perceiver as to the source of his perceptual beliefs, and his consciousness of their source in

sensation is itself the ground of his confidence. Everything epistemologically relevant is perspicuous to him, which is not to say that error is impossible.

Consider the common notion of ‘extra-sensory’ perception. Such perception must be supposed, like all cognition, to involve mechanisms linking what is known to the subject’s beliefs about it. Their peculiarities lie rather in this: the subject is supposed, for example, to ‘feel’ that someone is watching him from behind, or that the next card turned up will be an ace, whether by the pricking of his thumbs or independently of any accompanying sensations, but in either case without appropriate sensations, ‘fitting’ the state of affairs perceived. Thus a typical and even unique sensation may arise, it may be a part of the causality of the belief, and it may be recognized as such by the subject. But still, if that sensation is nothing but a quite peculiar tingling at the back of the neck or a blurring of vision or any sensation or aspect of a sensation which presents something other than the state of affairs in question, and so is not *appropriate* to the belief which it causes, then the faculty is not a complete sense but a mere belief-producing mechanism the operation of which contingently involves sensation. That is not to say that such a faculty, with or without ‘characteristic’ sensations, could not give rise to knowledge. But it would not be immediate or primary knowledge. If past experience (or for that matter detailed knowledge of the physiological mechanisms) taught us that such ‘feelings’ or conscious inclinations to believe are in general to be trusted, they would still arise for the subject as mere hunches, of proven reliability perhaps, but lacking the immediate and evident credentials of what is grounded on sensation, of what we evidently see or hear or feel to be the case; and so can know immediately. The subjective mysteriousness of such postulated secondary knowledge is no accidental characteristic of it, but the mark of its dependence on perspicuous first-order knowledge. It is the contrary of ‘evidence’. ‘Extra-sensory perception’, if it exists, is therefore not really or fully perception. But even if we grant the name ‘perception’, the very possibility of a distinction between sense-perception and extra-sensory perception depends on the fact that in sensation some things are presented to consciousness and others are not. The question as to which things are presented is independent of, or prior to, the question as to which beliefs those sensations may help to cause. Only that priority, i.e. the *intrinsic* intentionality of sensation, can explain the essential epistemological role of sense-experience or make it seem anything but an accident that sensations occur in the process which gives rise to perceptual knowledge. It is the ‘evidence’ of the senses which explains their independent authority, and which vindicates the view that it is in their deliverances that knowledge has its foundations. The notion of the perspicuity and evidence of perceptual knowledge will be further elucidated in [chapter 21](#) below.

Does perceptual knowledge have a firm boundary?

Any account which ascribes an independent authority to the deliverances of sense, as the foundations of our knowledge of the world, needs to say something fairly definite about the limits of that authority and the scope of perceptual knowledge. How far that is possible must be one of the chief deciding factors in the dispute between ‘foundationists’ and ‘anti-foundationists’. Not only do anti-foundationists commonly appeal to the alleged absence of any boundary to perceptual knowledge which is not relative to the perceiver’s wider knowledge or ‘theories’, but if it could be established that there is such a boundary, there would be at least *prima facie* reason for foundationism.

In favour of the latter there is the strong intuitive plausibility of a distinction between what we actually do or can perceive with the senses and what we may say that we perceive, but do not literally perceive. We can readily accept Descartes’ point that we may say that we see people from a window, when all we really see are their hats and coats. The antique-dealer who claims to see and feel that a chair is genuine Chippendale really sees or feels no more than its shape, the crispness of its carving, its patina, the structure of its joints and so forth. It seems beyond dispute that the attribute of having been made in a certain workshop more than two hundred years ago could not possibly be, in literal truth, an object of sight or touch. Yet in other cases such simple appeals to common sense may falter. A *ball* or a *coin* is an object made or employed for a specific purpose, a *goat*, as a member of a natural species, has a familial relationship to other goats, a *child* is a human being below a certain age; and yet it is plausible that we can literally see that an object before us is a ball, a coin, a goat or a child.

Here it may be tempting to fall back on a traditional model according to which sensation and perception stand in a quite other, more complex relationship than the one proposed above. The general model has its roots in Cartesian theory but was given its classic statement by Kant, whose philosophy hinged on the thought that we employ ‘concepts’ to ‘work up the raw material of sense-impressions into that knowledge of objects which is called experience’.²⁸ On such a view, what we perceive is the product of the ‘conceptualization’ or interpretation of raw sensory input, and it is the process of interpretation which generates intentionality. Crudely, we have a visual experience of a coin, or see a coin,

when it is through our concept *coin* that the sensory input has been processed. Consequently, as the recent advocate of a distinction between ‘sensational’ and ‘representational’ properties of experience has put it, ‘no one can have an experience with a given representational content unless he possesses the concepts from which that content is built up’.²⁹ The question now arises how, if at all, we are to draw a distinction between concepts which can be employed to build up perceptual experience, and concepts, such as the concept of a genuine antique, which seem to take us beyond the bounds of possible perceptual knowledge and into the realm of what we must infer from perceptual knowledge. Is it perhaps that we can identify a class of concepts linked in some peculiarly tight way to the sensational properties of experience, and which therefore define the representational properties of experience in itself, prior to inference?

Anti-foundationists reject such an enterprise, and are likely to reject the notion of a ‘sensational’ component of experience altogether. Their employment of the Kantian model will masquerade as a critique of it. There is no ‘raw material of sense-impressions’, and so nothing to serve as an anchor for supposedly experiential or observational concepts. Richard Rorty explains the position with impressive brutality:

There is, to be sure, a place for the notion of ‘direct knowledge’. This is simply knowledge which is had without its possessor having gone through any conscious inference. But there is no suggestion that some entities are especially well-suited to be known in this way. What we know inferentially is a matter of what we happen to be familiar with. Some people (those who sit in front of cloud-chambers) are familiar with, and make non-inferential reports of, elementary particles. Others are familiar with diseases of trees, and can report ‘another case of Dutch elm disease’ without performing any inferences.³⁰

To offer such a purely negative definition of ‘direct knowledge’ is in effect to explain perceptual belief in purely causal terms, as the first belief to be caused by the process of sensation in the circumstances of whatever other antecedently held beliefs. Perception is run together with perceptual belief, which is *grounded* on nothing, unless on its unconscious relation to all antecedent beliefs.

Not all anti-foundationists make the boundary of what can be directly observed to be the case relative in just this way to the antecedent knowledge of the observer. Quine, for example, identifies a class of ‘observation sentences’ as those ‘occasion sentences’ the ‘stimulus meanings’ of which vary very little or not at all ‘under the influence of collateral information’: i.e. assent to, or dissent from, such sentences in the presence of the same sensory stimulation would be more or less constant and unaffected by the speaker’s past sensory stimulation. ‘Bachelor’ is therefore not observational, or not as observational as ‘Rabbit’. But the latter is in turn less observational than ‘Red’, since, for example, a difference in their knowledge of the local fauna might lead the same glimpse to stimulate

one person's assent to 'Rabbit?' but not another's. Yet even 'Red' is said to be imperfectly observational. Information about unusual lighting might restrain one person's assent to 'Red?' which another freely gives. On the other hand, 'observationality' is said to be relative to the temporal 'modulus of stimulation'. As we make the 'modulus' longer, differences among speakers of the language with respect to the 'stimulus meaning' of any 'occasion sentence' will tend to be ironed out, for collateral information will come to be regarded as an earlier part of the stimulus itself. If we make the 'modulus' long enough, so we are told, even 'Bachelor' will tend to become observational.³¹

By rejecting the intuitive notion of appearance, or of what is evident to the senses, Quine removes the possibility of appealing, in explanation of the extremely variable stimulus-meaning of the occasion sentence 'Bachelor', to the obvious point that bachelors have in the nature of the case no characteristic appearance. Indeed, by the kind of paradox endemic to behaviourism, the latter platitude has to be 'explained' by the former technicality. The differences between such fundamentally different sorts of predicate as 'red', 'rabbit' and 'bachelor' are consequently treated as differences of degree. Most importantly, no distinction is drawn between the way in which 'collateral information' can defeat an appearance or deliverance of sense (such as the appearance of an object as red) from the way in which it can help to determine a judgement about a perceptual object which goes *beyond* appearance. There is also a third way, undistinguished by Quine, in which background knowledge can influence perceptual belief, i.e. through its influence on the way things appear to the senses themselves. Past experience of five-pound notes may lead us actually to see a discarded wrapping differently, as a five-pound note. Or perhaps only after being told that a picture is of an old woman do we have a certain visual impression of it. These natural, in themselves unproblematic, distinctions are smothered under Quine's artificial continuum.

The claim that the limits of the observable are relative to observers' theories, or are a matter of degree, goes hand in hand with the claim, popular among philosophers of science, that there is no hard distinction between experience and theory since the terms in which observations are reported will always be to some degree theoretical. A common presupposition here is that any account of a state of affairs with causal implications is theoretical, so that even the report that there is something red in front of me counts as theoretical. Since anything with objective existence has effects, every existential proposition is on this account theoretical. Yet such a conception of the theoretical is unhelpfully broad. Even statements like 'The old cat is sitting on the woven mat' are best regarded as pre-theoretical, although no one could understand its terms who did not possess some sort of structured apprehension both of the natural world and of human practices. Not all structure is theoretical. Our natural view of ourselves as material objects in spatial and causal relations with other such objects is not a sort of folk theory about the way things are which is open to replacement by the views of science as one scientific theory is open to replacement by another. It is intrinsic to

perceptual knowledge, involved in the very content of sensation. An appropriate response to the paradox that every report of an observation is framed in the terms of some theory is to appeal to something much more like Locke's distinction between 'natural history' and 'science' than most present-day philosophers of science would allow us. A sign that such terms as 'gold' or 'man' are pre-theoretical is the possibility of their being precise translations of terms in, say, ancient Chinese or any other language spoken by people with vastly different theories about gold and men from those of most present-day English-speakers. A statement's being pre-theoretical, however, does not entail that we can have perceptual knowledge of its truth: for example, we cannot literally see a thing's kind or age as we can see its shape and colour. The question where perceptual knowledge ends and the question where theory begins may be connected, but they are not the same question. There are here two lines of division.

Another issue remains. Does the sense-experience on which perceptual belief and knowledge must be grounded fall into two components, blind 'sensation' on the one hand and the 'concepts' to which the experience owes its intentionality on the other? What makes this analysis of experience both plausible and yet useless is the slipperiness of the notion of a 'concept'. Peacocke, for example, who has advanced the principle that 'only those with the concept of a sphere can have an experience as of a sphere in front of them,' gives two accounts of what a 'concept' is. First, it is 'a way of thinking of a property': yet, since Peacocke uses the term 'thinking' in a broad sense to include the having of experiences, that explanation would make his principle an uninteresting tautology. His other explanation, that a concept is 'a mode of presentation of an object', achieves the same result: for modes of presentation may, he says, be 'perceptual'.³² We hardly need to be told that only those capable of experiencing or perceiving a sphere can have an experience as of a sphere in front of them.

We could avoid that kind of vacuity, of course, by adopting a less thin-blooded, in some ways more traditional, notion of a 'concept', taking the concept of (say) a trapezium to be what must be grasped in order to grasp the *meaning* of the term 'trapezium' and its synonyms. Possession of the concept is on that account possession of a piece of classification. The difficulty now is to show that possession of the concept of a trapezium is a necessary component of the capacity to see trapeziums, or to see things as trapeziums. It is, of course, true that we need to know the meaning of 'trapezium' in order to respond to the question 'What did you see?' or 'What did it look like to you?' with the answer 'A trapezium'. Yet another, linguistically less direct way of answering such questions is available, and must be employed where language is deficient: we may reply, 'It looked like this', while drawing or pointing to a shape resembling what was seen. The latter response tells just as well as the former (and in some ways better) *how* the shape was seen. To make it, no concept is required.

Hence there is no reason why two people, only one of whom possesses the 'concept' (in this sense) of a trapezium, should not both have just the same visual experience in which a trapezium is presented. Both see the figure in the

same way, and see it as it is. If it is appropriate to say of one of them that he is seeing it as a trapezium, then it is appropriate to say it of both. And it is appropriate to say it of both because of the possibility of misperception: a third observer might see the trapezium distortedly as if it had no sides parallel. It should not, then, be supposed that our coming to recognize and classify a shape as a trapezium (any more than our recognizing it as a shape seen before) essentially involves some alteration at the level of perceptual experience. For what would that alteration be, in the case of a trapezium? Suppose that recognition did in fact cause an aspect-shift, so that we came to see one of the parallel sides as the base. A trapezium does not look more like a trapezium (although it might look more like a table) after than before such a shift; and in any case just that aspect-shift might happen *without* our recognizing the shape as a trapezium.

Aspect-change and misperception supply contexts in which the expression 'see...as' does much of its work, and they very well illustrate the worthlessness of the notion of 'concepts' to a philosophical account of what perception is. In a familiar example of aspect-shift, a diagram can be seen either as a white Maltese cross on a black background, or as four black shapes of peculiar proportions in quasi-dynamic relation. We do not need the concept of a Maltese cross in order to see the diagram as a Maltese cross, any more than we need a special concept unmarked in any language in order to see it the other way. If it is insisted that a different *concept* corresponds to each *aspect*, that tells us more about the present vacuous use of 'concept' than about perception or 'seeing as'.

'Concept' is one of those particularly dangerous philosophical terms ('induction' is another) which is commonly stretched between theory-neutral and deeply problematic senses. If 'induction' is defined as whatever kind (or kinds) of argument gives rational support to predictions, hypotheses and the like, then no doubt our rational predictions are all inductively supported: but if it is defined as argument the logical form of which is given by some hopefully patched-up version of (say) Russell's 'principle of induction', then we had better say that rational predictions are never inductively grounded, or only ever in part and in special contexts. Similarly if 'concepts' are stipulated to be whatever is necessary for thought and experience of objects, then of course we cannot think or perceive without 'concepts'. Yet if the notion of a concept is allowed to have some backbone by retaining its traditional link with 'terms' and determinate logical form, then the principle that we cannot think or perceive without concepts commits us to one questionable thesis or another: either, as Locke assumed, that thought and experience prior to language possess determinate logical form in their own right; or that perception and thought are impossible without language, so that saying comes before seeing. Both the ancient error and the modern paradox can be avoided, if it is recognized that we do not after all need to possess the concept of a trapezium in order to see one.

When verbs of perception are followed by a prepositional object, the connection between what we can see and what we can say may seem an

especially close one. For it may seem that, unless we know what a 'trapezium' is, we cannot see *that* a figure is a trapezium. This idiom brings perception close indeed to belief. Because we normally trust our senses, to say without qualification that X sees that *p* carries the conversational (but not logical) implication that X believes that *p*. As it has also been remarked above, in order to ascribe a belief to a being without language it is necessary to express it, to give it logical form, and so in a way to distort it. In general, we may be reluctant to attribute propositional attitudes in terms which the subject does not possess or, worse still, in terms of a classification based on differences and resemblances of which the subject could not be aware, or in which it could not be supposed to have an interest. Yet such reluctance as we may have to express a dog's belief as the belief that its master is about to take it for a walk is not absolute, despite the dog's ignorance of what a 'master' is, and its lack of the concept *walk* with which to single out parts of its master's and its own activity. The dog evidently has a certain expectation which, if we are to take note of it, we must perforce express *somehow*. That might be called 'the Problem of Expression'. Our reluctance to say of someone who lacks the concept of a trapezium (or who does not *recognize* the figure as a trapezium) that he sees *that the figure before him is a trapezium* is perhaps all the stronger because there is a ready alternative with the same import: we can say that he sees it as a trapezium, or as it is. But whatever we may prefer to say, the nature and content of such a person's sensation, the end-product of the process of perception, is not essentially affected by his possession or lack of the concept of a trapezium. The sensation or sense-experience could be the same in either circumstance. In either case the sensation lacks logical form, despite its intentional content. Intentionality is more primitive than linguistic concepts.

The notion of the conceptualization of raw data (if not the currently more fashionable notion that no data are raw, that all come conceptualized) has perhaps always drawn support from seemingly plausible speculations about the experience of babies, in or just out of the womb. It is a specious view that in human development initially pure, primitive and objectless sensation acquires intentional content through being ordered or interpreted in accordance with learnt concepts or constructed theory. Yet such an explanation of an infant's fairly imprecise relationship with the world is more mysterious than what it purports to explain. No doubt the content of our first sensations is relatively indeterminate, as the content of our sensations on first waking may be, but it is difficult to see how they could lead to anything by way of 'concepts' or knowledge unless they included some apprehension of things in space and of their action on us. To suggest that a baby's pain is not felt in its body is to make both its pain and its subsequent acquisition of a 'concept' of its body more mysterious, not less so. Among the first indications of an infant's awareness are those of its awareness of things in spatial and causal relations to itself. It may be proposed that it has devoted the very brief period between its birth and its reaching out for things to formulating a 'theory' of objects in space on the basis of the constancy and

coherence of blank data. But that is to engage in speculation more philosophical than entirely sensible.

The scope of perceptual knowledge

Lockean 'sensitive knowledge' is pre-theoretical and immediate, but extends less far even than the primitive awareness of objects which we need to ascribe to a baby. It does not include knowledge that the world is really and actually spatial. That 'a piece of Manna of sensible Bulk' and in motion can produce in us an idea of motion which 'represents it, as it really is in the Manna moving' may be something 'every Body is ready to agree to',³³ but on Locke's official account it is not something evident to the senses or arrived at without speculative thought. He did accept the metaphysical principle that whatever exists and acts is particular and therefore has position in objective space and time, but even that goes beyond the deliverances of the senses.³⁴ His argument implies that the world as immediately known through sense-experience is entirely a world of powers. It was precisely this doctrine which was soon to seem to Berkeley an open invitation to advance a rival thesis about the reality to which those powers belong, i.e. that it is non-spatial and spiritual.³⁵ In order to do justice to the immediacy of our knowledge that the world we perceive is in itself spatial, we need something Locke does not give us: an argument which implicates things' being spatial, and being perceived as spatial, with the very possibility of perceptual knowledge of them.

Such an argument seems to flow from a certain defect in Locke's official position. He treated sensory effects or ideas, for the purposes of his theory, as 'simple and unmixed', but as perceived to 'coexist'. Yet the coexistence of qualities in an object cannot just be a matter of the co-occurrence of their effects in the mind: the one is neither a necessary nor a sufficient condition of the other. We characteristically perceive the 'coexistence' of sensible qualities as coexistence at a place, or rather as features of some space occupant located through its spatial relations to ourselves. Not that every perceived sensible quality is always immediately perceived as being at the location of its possessor. We characteristically associate a smell with the odoriferous object either by estimates of the smell's³⁶ intensity at different places, or by exclusion, i.e. by removing objects in turn. The two methods are in principle the same, and both (the first only more obviously than the second) rely on a smell's having a

perceptible location in its own right, as well as on our ability to locate its substantial cause, the possessor of a 'smell' in a slightly different sense.

Because Locke left our immediate perceptual knowledge of space out of account in his official theory of sensitive knowledge, however much he may have assumed it in off-duty moments, his treatment of our immediate awareness of causality is also defective. For he limited the latter, in effect, to awareness of the involuntary character of 'actual sensation'; whereas in sense-experience we are also characteristically aware of how things act on us, in that we are to some extent aware of perception itself as a physical process. Not that we perceive the mechanics of perception in any detail, of course, but, in being aware of the spatial relationship between the object and the organ of perception, we are characteristically also aware of their causal relationship. The issue might be approached through the question whether certain sensations of touch could ever be (as has so far been taken for granted) 'accidentally' veridical. Suppose that entirely because of some damage to my spine I have the sensation of something pressing painfully into my leg, which is in fact anaesthetized. If by chance there actually is something pressing there, it does not follow that my sensory state is veridical. First, although something is pressing into me, nothing is actively hurting me in the way in which I *feel* something to be hurting me. Arguably some such sense of causality is necessarily involved in all perception. In general, in being sensorily aware of objects, we are aware of their striking the senses. In that case no sensory state could ever be *in toto* accidentally veridical.

Analogous considerations can cast light on the most important of Locke's 'ideas of one sense', the only one to be accorded the status of primary quality, or a chapter to itself, in the *Essay*. The role of the idea of 'solidity' within his system is not all obscure, for in it he saw the best hope of the theory of atoms and the void to provide a genuine explanation of the sensible world. Solidity is presented as the actual or intrinsic property of matter which, *contra* Descartes, distinguishes it from empty space and gives rise to its powers of impenetrability, 'mutual impulse, resistance and Protrusion'.³⁷ The necessary connection between solidity and 'receiving or communicating Motion by Impulse', Locke felt, is at least partly intelligible.³⁸ Solidity in itself is a property of matter like extension and motion, perspicuous because like them it is represented by an idea which resembles it. According to the doctrine of simple ideas, more neither can nor need be said: 'If anyone ask me, what this Solidity is, I send him to his senses to inform him: Let him put a Flint, or a Foot-ball between his Hands; and then endeavour to join them, and he will know.' The claim was, then, that the simple idea of solidity is an indefinable *quale* which represents and 'resembles' that property of matter in virtue of which objects press upon, and resist the pressure of, other objects including ourselves. The whole model crumbled before the onslaught of Hume, but Hume characteristically rejected, not the deeply suspect notion of the simple sensory *quale*, but the possibility that this item, a tactual feeling, should be *like* solidity, if solidity is that in the object which constitutes its impenetrability.³⁹ The alternative path, just as destructive and more

illuminating, is to find an acceptable sense in which our sensations of solidity do represent their object as it is, but to reject the conception of the simple, blank *datum*.

Solidity is after all a sensible quality, but it is so, not because sensations of solidity are the simple *qualia* of Locke's theory, but because they are sensations of a kind of causality. It is significant that the 'idea of solidity', as it is introduced in the *Essay*, seems to comprehend all sensations of pressure or resistance, even the virtually ever-present feeling of 'something under us, that supports us, and hinders our farther sinking downwards'. Yet he was doctrinally bound to regard these descriptions of their objects as merely indirect accounts of such sensations: as if they identify the common sensation or idea of solidity by reference, not to its intrinsic nature or content (which, he supposed, can be given only in actual sensation), but to conditions under which we have that idea, conditions themselves mere consequences of the solidity of objects. That is what is wrong with Locke's model. What is common to all sensations of pressure, resistance, support, weight and so forth is just that that is what they are. The various ways in which we are tactually aware of our mechanical interaction with objects which offer resistance to our bodies are not bound together by some simple, inexpressible common 'quality'. They can be grouped together just *because* such mechanical interaction enters into the intentional content of the sensations involved in them. No particular feeling of contact or pressure is distinctly conceivable, as both Locke and Hume presuppose, in abstraction from our general sensory awareness of the interaction between our bodies and things in space. It is because of all this that, although there are characteristic ways in which solid objects feel, depending on the circumstances, the perception of solidity is different from that of colour and smell. For there is nothing sense-relative about mechanical interaction and yet it can be directly perceived: i.e. it can enter into the content of our sensory states. Yet to say that is hardly to endorse Locke's motive for classifying solidity as a primary quality. There is no hope for his view that there are simple and qualitatively uniform experiences of solidity which are peculiarly capable of giving a sense to a fundamental term in his favourite physical theory.

The general point that the spatio-causal relations between the perceiver and the objects or states of affairs perceived are themselves the object of sensory awareness casts an important additional light on the perspicuity or 'evidence' of what I have called 'primary knowledge'. Primary knowledge is perspicuous in that one who has it knows *how* he knows what he knows: its ground and *ipso facto* its causality, in broad terms, are perspicuous. Any knowledge of the world derives from the state of affairs which is the object of that knowledge, and primary knowledge is knowledge grounded on something which displays that derivation. When I see a skua diving at my head or feel water splashing on me, the content of my sensation includes not only the skua or the water but its spatio-causal relation to me. I am aware of how I know of the presence of such an object in so far as I am aware, not only of the source in sensation of my belief, but also

of the way in which the object is acting on my senses—which does not of course mean that I am aware in detail of the mechanisms involved. Analogously, my reasons for a non-perceptual belief which constitutes primary knowledge will include an understanding of how the relevant state of affairs has come to impinge on me. Roughly, my reasons for the belief will mark out a causal route from state of affairs to belief. In secondary knowledge, whether quasi-perceptual (as in the case of ‘blindsight’) or non-perceptual (as with any item of traditional knowledge the origin of which has been entirely forgotten), the derivation of knowledge from its object is not in the same way perspicuous.

Perceptual knowledge is, then, knowledge of spatial and causal relations between things and ourselves. But it is more than that. The things related must impinge on our sensibility in otherwise distinguishable ways, feeling hot or cold, looking red or white, tasting sweet or sour and the like. There could not be a perceptual object which lacked ‘secondary’ qualities. Yet unless we had an immediate and authoritative apprehension of spatially extended, spatially related things, those aspects of our sensation which are the basis of our attributions of secondary qualities simply could not sustain the role. Crudely, we could not possibly have perceived only the secondary qualities of things, because we need a spatial framework on which to hang them: we need perceptual knowledge of things-in-space on which to hang both speculation and confessions of ignorance about the real, underlying nature of the observable, non-spatial differences between these things.

Malebranche tried to capture the relationship between the perception of spatial attributes and the perception of secondary qualities by means of his distinction between *idées* and *sentiments*, perception and sensation. For all that has been said above against such a distinction and in favour of understanding perception simply as successful sensation, it may be helpful to set it against Locke’s notion of ‘signification’. Locke put ideas of secondary qualities on a par with ideas of primary qualities just because he did not want their truth, adequacy and reality to be in any way impugned. ‘Resemblance’ between ideas of primary qualities and their causes is a luxury we are lucky to have, if we have it. So far as Locke’s theory goes, we might only have had ideas of secondary qualities, and yet still have had perceptual knowledge: or like angels we might only have had ideas which resembled their causes. For Malebranche, on the other hand, human beings perceive clear *ideas* of the modes of extension, but need to pick them out as particulars by means of *sensations*. Both are necessary for the perception of particulars.⁴⁰

Reid advanced a view rather similar to that of Malebranche. For him, sensations merely accompany our ‘direct and distinct’ perception of primary qualities, whereas our perception of secondary qualities is ‘indirect and obscure’: we conceive of them simply as the causes of our sensations. For both these philosophers it is, or should be, apparent to us on the briefest reflection which sort of quality is which. As Reid put it, the ‘distinctness of our notions of primary qualities prevents all questions and disputes about their nature’. Leibniz, who in

effect reduced the status of spatial attributes to that of secondary qualities, in Reid's view overturned the whole evidence and authority of the senses. That seems right. Yet Reid seems to have been less sensitive than Malebranche to the interdependence of the perception of primary and of secondary qualities. 'Sensations', on Reid's account, are only contingently associated with perception at all, and if he comes close to ruling it an impossibility that all our perceptions might have been 'indirect and obscure', he neither offers to explain why it is impossible nor recognizes the equal impossibility that they should all have been 'direct and distinct'.⁴¹

The chief failing of both Malebranche and Reid on this question, however, is that, in separating 'perception' and 'sensation' as different sorts of mental occurrence, they made the connection between them unintelligible. As much as Locke did with simple ideas, they presented 'sensations' as blank data. Yet if we talk of 'the sensation of yellow', we are talking of what is nothing but an abstraction from the sensory state, from the state of sensory awareness as a whole. It is not something we could conceivably have *except* as an aspect of our sensory awareness of things, the appearance to us of space and of things in space. Equally, as Berkeley recognized, the appearance of secondary or sensory attributes is essential to the appearance of spatial qualities.

Reid's argument, like many modern arguments in the area,⁴² relied heavily on the example of physical pain, which he treated as a 'sensation' contingently coupled with the 'perception' of a disorder in a part of the body. The location of pain, he thought, has nothing intrinsically to do with the 'sensation', but is the 'perceived' location of the disorder of which we have an 'indirect and obscure' conception as the unknown cause of the pain. More will be said about pain in [chapter 23](#) below, but Reid's account is as unconvincing as it would be to claim that the location of colour is something separately perceived which does not in any way enter into the sensation of colour, as if the sensation could in principle exist in us in the absence of any apparent extension. To have or feel a pain in the left foot is to have a certain sort of sensory awareness of one's left foot, which may or may not be disordered. Like the word 'feeling', the word 'pain' may here be used either for the state of awareness itself, or for a sort of secondary quality of the foot, for what is felt there. As the second sense indicates, 'pain' in the first sense is integrated with the sufferer's normal, less striking perceptual awareness of the body. Reid's mistake comes down to supposing that its painfulness, as abstracted from the whole perceptual state, exists on its own as a mental occurrence separate and different in kind from that part of the state which is the bearer of intentionality and the determinant of the place of pain. That it is a mistake follows from the necessity of the principle that physical pain must have location, as colour must be extended.

The same principle applies to sensations of heat and cold, although here the situation is further complicated. To feel heat can either be to feel it in what is touched, or, as in feeling pain, to feel it in a part of the body. Often the two go together: the poker feels warm and my hand feels warmed by the poker. Such a

case is not the accidental conjunction of distinct sensory data, any more than the tactile awareness of external objects is subjectively quite distinct from the awareness of one's own body. The latter itself involves more than one mechanism and more than one subjective ingredient. Sensations of touch, of heat and cold, 'proprioceptive' sensations (sensations of one's body as an object extended in space), pain, 'vestibular' sensations (sensations of balance) are in normal circumstances closely harmonized, and in some cases what each contributes to our awareness of objects and ourselves in relation may only be identified and isolated through damage and loss. That is presumably why several forms of sensation involving distinct mechanisms have been liable to be attributed to a single sense of 'touch'.

The senses of taste and smell, on the other hand, have generally been supposed to have proper objects uncontaminated by the spatial, and it might be thought that they supply easy counter-examples to the thesis that the intentional objects of sensation are necessarily things in space. Yet these supposedly distinct senses are in fact very like the capacity to feel heat and cold, except that the parts of the body in which tastes and smells are experienced—the parts, that is to say, awareness of which is involved in tasting and smelling—are confined to very specific areas of sensitivity. We taste an object in the mouth, and the awareness of its taste is as integrated with the tactile awareness of its presence there as is the awareness of its heat. We can experience a taste without feeling a solid object in our mouth, of course, but the taste is still experienced as in the mouth. In the same way we smell a smell in the nostrils and palate. Something with a strong, acrid taste or smell, like something which is too hot, can cause a peculiar sensation of pain in the sensitive area which may blot out the sense of an object external to the body as we become more intensely, unpleasantly aware of the sense-organ itself. But some awareness of the general location of the organ is involved in all smelling and tasting.

We can perhaps imagine what it would be like to have a directional sense of smell (if human beings are in fact totally without one) by analogy with the sense of hearing. For although we do not hear shapes, the spatial notions of direction and distance seem as necessary to an account of the immediate content of hearing as they are in the case of the considerably more accurate sense of sight. The phenomenological characterization of a sound as coming from the left is incapable of translation into purely qualitative, non-spatial terms. Otherwise it would be impossible for qualitatively identical sounds to come from different perceived directions, as they evidently can. Because of the directionality of sound we can hear movement, not merely in the sense that we hear sounds caused by movement, but in actually hearing something pass by. Moreover, if the same sound, such as 'distant' shouts in a radio broadcast, can be heard either as distant shouts or as the nearby product of a radio, we have an auditory phenomenon comparable to the well-known phenomenon associated with three-dimensional vision, when we can see the same drawing of a transparent cube either two-dimensionally or three-dimensionally; and then either with one side at the front or the other. Yet it

seems that we can only make sense of such a possibility if the spatial 'field' of hearing is taken to be integrated with that of our non-auditory awareness of the body. What we hear as very close is heard in relation to the head, but the head is not itself heard. What is heard as coming from a certain direction is heard in relation to one's unheard body. Another illustration of this integration is perhaps supplied by the auditory/visual illusion upon which ventriloquism relies. Even if the sound is not coming from the moving mouth, we can hear it as if it is. We have a much better perception of its actual direction with our eyes shut. There are not, therefore, several sets of apparent directions, auditory, tactual, visual, proprioceptive, vestibular and perhaps olfactory, which we learn, or are innately inspired, to associate or identify with one another. There is one space of which we can be aware in different, but essentially integrated ways. The five (or more) senses are not distinct inlets for quite disparate, internally unconnected data.

It may help to elucidate the notion of an integrated sensory field if it is seen as an element in an argument that the scope of perceptual knowledge is in a particular respect much wider than it has traditionally been assumed. My claim is that a judgement which links the objects of different senses may itself be, and very often is, an immediate *perceptual* judgement, directly grounded on the deliverance of sense. Thus it is not normally as a result of inference, habitual association or the like (although in a few peculiar cases of disorientation some sort of inference may be to the point) that I *judge* the object I feel with my hand to be the object I see. Quite simply, I *perceive* it as the same: the identity enters into the intentional content of sensation, and of my total sensory 'field'. In general, too, we are aware of the place or point of view from which we see things. Our visual awareness of things in space, including what we can see of our own body, involves some awareness of the relation which they bear to the part of us from which we see. But this awareness is integrated with, indeed involves, our general tactile and proprioceptive awareness of the head and its relation to the rest of the body. No experienced correlation is needed to link a dazzling light with a pain in the eyes, as it would be if, by some quirk of neurophysiology, such a light caused a pain in the feet. These examples are analogous to our immediate (if imprecise) ability to point towards a source of sound, an ability which is based on the heard direction of the sound and which essentially involves the integration of our hearing with other senses, in the first instance, no doubt, with our proprioceptive 'body model', and so with our powers of voluntary movement. For the senses are not only integrated with one another, but with our capacity to act in the world. If we see something hurtling straight towards our point of view, we see it as coming towards the unseen front of the head, which we instinctively duck or protect. Oliver Sacks's graphic account of 'the disembodied lady' illustrates the special and unsurprising importance of the proprioceptive sense for bodily action, for the loss of it was, as in similar cases, immediately incapacitating. A sort of effective recovery was achieved, however, precisely because the sufferer found it possible to move and control those parts of the body she could see. There was

also only a slight impairment of 'light touch', so that the wind on her skin gave her some sense of her body.⁴³

If all this is at all near the truth, it follows that there can in principle be sensory illusions like the standard optical illusions, but involving intersensory relationships. One example, I have suggested, is supplied by ventriloquism. Another, which may not strictly be an illusion but is at any rate a form of intersensory misperception, is supplied by the case of prismatic spectacles which invert the visual field. Since everything seen is (until the brain adjusts) seen as inverted, the frame of reference for the apparent inversion must lie outside the deliverances of sight, in the sensory field as a whole. That is to say, the frame must be supplied by other senses (in particular, presumably, by one's sense of one's body) in so far as they are integrated with vision. Berkeley explained such total inversion simply as a judgement resulting from the reversal of the normally constant conjunctions between felt eye-movements and such objects as the 'visual' sky and earth. Yet that kind of explanation presupposes what is palpably false: that the inversion is not phenomenal or aesthetic at all, but relates only to inferences from one sense to the other, an object of judgement rather than of immediate sensation. The first explicandum here for the philosophy of perception is how such a *phenomenon* is possible, and Berkeley in effect says that it is not.⁴⁴

Unless the principle of integration is true, the relationship between sight and touch could consist in nothing more than that the same object can appear in the same way (e.g. as a cube, or as a globe) to both senses. It could not be the case that the object seen appears to be numerically the same object as the object felt (*or* that they appear *not* to be the same object). A thing could not look to be in the same place as it is felt to be; nor could it, by some refraction of light or the like, look as if it were in a different place from the place in which it is felt to be. The notion of a unified sense-field is therefore needed to solve the traditional problems raised by 'ideas of divers senses' and the perception of space. We need to understand that I can have immediate perceptual knowledge that the foot the shape and the size of which I feel with my hand, the foot which I see, the foot the coldness of which I feel both directly and with my hand, and the foot which is the object of other direct bodily sensations including pain, are one and the same foot. As it has been argued above, the objects of other senses than vision and feeling are incorporated into the same sense-field. If sight and touch are peculiarly alike, it is at a more specific level: for example, we can perceive the *shapes* of things outside us by both sight and touch, but not by other senses.

Historically (and naturally enough) the philosophical fragmentation of the sense-field went hand in hand with the fragmentation of its objects into clusters of sensible qualities merely associated in our experience. Both, as we have seen, were involved in the seventeenth-century arguments that the truth about the nature of things is not delivered by the senses. The famous reasoning of Descartes' Second Meditation was designed to show that the unity of the thing itself lies behind the sensory manifold, to be grasped only by the intellect. Locke's theory of sensitive knowledge and of the truth of simple ideas did not

diminish the tendency to disintegration. Although he himself maintained the notion of a unity available to the senses in the perception of 'coexistence', that notion is too impoverished to be satisfying or even coherent. Much philosophy since Locke's time has been devoted in one way or another to the problem of what can be supposed to unite disparate sensible qualities in the mind. Is the experienced unity of physical objects to be attributed to principles of the imagination, to Kantian categories, to a quasi-scientific theory or to the structure of language? All such answers presuppose that some unifying and discriminatory scheme is brought to the chaotic diversity of raw sensory input in order to generate experience of objects, whether that scheme is regarded as a merely natural, or as a constitutively rational order of thought, as a universal feature of human nature or as a mutable product of human needs and purposes.

The present proposal is that physical objects can enter into the content of *sensation*, prior to our possession of any 'concept' or theory. Despite the variety of ways in which an object such as a rabbit or a lump of lead (or, *mutatis mutandis*, a quantity of stuff with contiguous parts) may impinge on us, the integration of the sensory field enables us to experience it as a unitary thing, as a dog no doubt experiences the rabbit it pursues and catches. There is therefore no problem, from the point of view of epistemology, as to how we travel from blank atoms of sensation to the unified objects of experience. If the latter had not been given in sensation, we could never have made the trip. If I see something red and square, it is easy to understand that there is no real problem about the unity of the redness and the squareness: and that any appearance of a problem arises from a misinterpretation of the possibility of abstracting the one from the other. We simply *see* that one and the same object is both red and square. So too we can simply see and feel that there is something red, three-dimensional, resistant and coherent, and therefore relatively enduring, in front of us. Perhaps it takes more time to become aware of a thing's material coherence than it does to see its redness. Perhaps the neurological process is more complex, as the brain has to put together more physical cues. But sensation is a process the intentional content of which is not restricted to the vanishing present. We do not, for example, (*pace* Hume) have continually to judge or to postulate or even to assume that its objects at one moment are identical with its objects at the immediately preceding moment. The continuity of physical objects is as evident to sense as their existence. The apprehension of the one is involved in the apprehension of the other.

A possible source of confusion arises from those examples of sensible material objects which naturally come to mind—coins, tennis-balls, cats and so forth. There may even seem no reason why we should not add bibles, bachelors and professors to the list. All are perceivable and all are physical objects. Yet the claim that we can have primitive perceptual knowledge that there is a physical object before us is not the claim that we can have such knowledge that there is a coin or a tennis-ball or, still less, a bachelor before us. Admittedly we often say that we see that there are such objects present, yet what makes an object a coin

or a bachelor neither is nor could have been visible. In the end we can appeal to the authority of the senses in support of the contention that a certain object is perceptibly just like a penny, but not that it is a penny. Counterfeit may be perceptually indiscriminable from currency. A traditional conclusion from this point has been that we do not see, but infer that the object before us is a penny. Why then does that conclusion have an air of paradox? Why do we talk so naturally not only of seeing pennies (which is understandable) but of seeing that objects are pennies, and even of seeing objects as pennies? The same question arises with respect to natural species. To identify the species of a living thing is to go beyond its sensible appearance, yet in characterizing the content of sensation nothing is more natural than to speak of 'seeing' pink rats, or of seeing a bush as a lion. On the other hand, the claim to see someone as a bachelor or as a professor would normally be taken metaphorically, to refer to some kind of judgement or imaginative performance rather than to sense-perception.

The explanation of such differences seems to be that, although the essential monetary role of a penny, like the essential status of a bachelor or a professor, cannot be given in sensation, that role is associated with a characteristic appearance as the status of a bachelor or professor is not. The determinate appearance of a penny may be a matter of mutable convention, but in the nature of the case pennies need as such some rather finely peculiar and enduring sensible qualities or other. The same goes for tennis-balls, although here function imposes narrower natural constraints on the dictates of convention. There is unlikely to be a fashion for square tennis-balls. In the case of species of living things the connection between membership of the class and appearance is entirely natural, and sensible appearance is commonly so characteristic of the species that it seems paradoxical to insist that we cannot just see that an animal is a rabbit and not a hare. Nevertheless the attribute of being a rabbit is not an object of sense as a rabbit or its colour is.

Is the traditional view therefore correct, that when, as we say, we simply *see* that something is a coin or a rabbit, in fact we *infer* that it is a coin or a rabbit from its appearance? Since there is no conscious inference in such habitual recognition, should we accept that there is unconscious inference? It may not be very clear what is at stake in this question, if not the boundary of perceptual knowledge, and of the immediately sensible, itself. Someone might argue that an appeal to the notion of unconscious inference is gratuitous: just as we talk of simply *feeling* that an object is hot or solid, when heat and solidity are each much more than a tendency of objects to feel a certain way to observers, so we ought to talk of simply *seeing* that an object is a coin, despite the complexities of what it is to be currency. As heat and solidity are attributes each with a characteristic feel, so a penny has a characteristic look. Yet such an argument ignores a crucial difference. If a cold object feels hot, then I conclude that there is something wrong with me, or something special in the circumstances, which will explain the illusion of heat. But if a clever counterfeit looks just like a coin, there is no sensory illusion to explain. (That is not, of course, to deny that an object may

occasionally look more like a coin than it really is by a kind of passing illusion.) It is possible to mark this distinction without recourse to the notion of unconscious inference or to any notion more recondite than that of recognition. When we recognize an object as a penny, our perceptual state may be unchanged or, if it is changed, the change follows, rather than constitutes the recognition. The recognition involves a sort of interpretation of what is given in sensation which goes beyond sensation. It is like the realization or understanding of how a mechanism is working which one onlooker may have and another lack, although both see the same events with equal clarity. Indeed, someone to whom the mechanism's parts are unfamiliar may see them more distinctly, rather than less so. Although not a constituent of our perceptual state, such recognition or realization or interpretation is a constituent of our state of consciousness. But that proposal looks beyond the boundary now under examination, and discussion of it is best postponed to [Part IV](#), below.

There is another question about the intentional content of sensation which also leads on to an argument in [Part IV](#). How, if at all, do particular individuals enter into the content of sensation, so as to be objects of perceptual knowledge? From one point of view, there seems to be no problem: everything we perceive is, of course, particular, and is perceived as particular. But from another point of view it can seem impossible that particulars should enter as objects into the *intrinsic* content of sensation, just because an essential condition of its reference to a particular object lies outside the sensory state, in the causal relation between sensation and object. In so far as we are concerned with intrinsic intentional content, we are concerned with general *types* of sensation. Thus two subjects in widely distant psychological laboratories may have visual sensations with just the same intrinsic content (and so of the same type), although they are looking at numerically distinct spinning discs. But if perceptual belief is belief which is grounded on sensation, and if that means that its intentional content derives from that of an antecedent sensory state, how can one of the subjects have a perceptual belief about *this* disc rather than the other disc? What in his sensory state picks out *this* disc in particular? A second, but related question concerns the perception of individuals otherwise known to us: how can we have perceptual knowledge that John has entered the room, or how can we see that we are standing by Magdalen Tower? For it is possible to imagine that we should have been in just the same sensory condition, without misperception, in the presence of a different man with the same appearance, or a different tower; and so with any individual.

A part of the answer to the first question is simply that reference indeed depends on a causal relation, as well as on a match between intentional state and object; but that perceptual belief derives its reference (if any) *as well as* its intrinsic content from the antecedent sensation. Yet there is more to it than that. Our sensations, as we have seen, are not bare representations which may or may not match the object which causes them. They present their objects as being in certain causal and spatial relations to ourselves and our sense organs. A particular thing is an object of our current sensations above all in virtue of its

satisfying just that spatio-causal constituent of their intrinsic content. It is related to us as it appears to be. In other words, in perception the extrinsic referential relation is itself perspicuous to us in virtue of the intrinsic content of sensation. The perspicuity of perceptual reference is, indeed, a part of the perspicuity of perceptual knowledge. It is the reason why, truistically enough, demonstrative reference has its primary role in the perceptual context. Moreover such perspicuous reference is the *only* sort of reference to belong to perceptual knowledge, strictly conceived. Strictly speaking, we cannot have perceptual knowledge that we are standing under Magdalen Tower. Yet all that means is that the evidence of our senses is not enough, since we also have to *recognize* this tower as Magdalen Tower. The problem now becomes one of explaining the reference, whether perspicuous or non-perspicuous, of thoughts other than perceptual beliefs. We must move on from an account of perceptual knowledge to an account of recognition.

Two modern approaches to sensation

Locke's theory of knowledge ascribes a central role to consciousness, but his causal theory of signification or representation tends, as we have seen, to evacuate experience and thought of the intrinsic intentional content traditionally ascribed to it. Very many modern philosophers would be inclined to deplore the former tendency and approve the latter, whereas I have done the reverse. Even those philosophers who have found it too difficult to pretend that they are not conscious have often taken their stand on the awfulness of pain, the subjective quality of colour and the like, rather than on the intrinsic intentionality of experience. The chief, and for some the only, argument for accepting the phenomenon of consciousness has been its obvious existence, rather than its indispensability for explaining the possibility of knowledge.

As one of the most notable of its current champions has remarked, 'Consciousness is what makes the mind-body problem really intractable.' It is not, he continues, 'analyzable in terms of any explanatory system of functional states, or intentional states, since these could be ascribed to robots or automata that behaved like people though they experienced nothing'.⁴⁵ Yet the intentionality ascribable to certain states of a robot lacking consciousness must be an extrinsic or relational attribute, in this, at least, like the intentionality of Locke's simple ideas or, more directly, Descartes' corporeal images. The present claim, on the other hand, is that we can give a philosophical explanation of knowledge in general and (as far as the argument has progressed) of perceptual knowledge in particular only by postulating states with *intrinsic* intentional content through which what is known is evident to us, the perspicuous source of our beliefs. Recognition of this role requires us to return to the traditional thesis that consciousness and intrinsic or primary intentionality go hand in hand. This reactionary contention (or piece of common sense, as you will) must labour upstream against two powerful currents of present-day philosophical thought, a certain sort of respect for science and a similar respect for language. It would only be courteous (and, perhaps, rewarding) to take some notice, as they press by in the opposite direction, of what these tendencies have to offer on the topic of perception and perceptual knowledge. But there will be no attempt here at detailed exegesis of the latest thing.

The philosophical respect for science which is in question takes the form of an insistence that we live in a world totally available to a certain sort of physical theory. Consciousness and intentionality, unless construed in a certain way, are liable to appear as awkward excrescences on the physical system. 'Functionalism' is a construal of them which removes the embarrassment while allowing a sense, if a thin one, in which we think, have experience and, indeed, can even be said to be 'conscious'. A combination of the earlier doctrines of behaviourism and physicalism, functionalism holds that psychological processes and mechanisms are physiological processes and mechanisms described in terms of their causal or functional relationships to other processes, and ultimately to sensory input and behavioural output. The strength of the model is hardly deniable. It lies not only in its ontological economy and stage-clearing for physical science, but in its accordance with ancient intuitions about existence and identity. What could the particular existence of mental events and processes be, if not existence in space? And what can occupancy of space be, unless a sort of physical existence? It seems to most of us only common sense, a rational presumption which any psychologist bears with him into the laboratory, that our mental functioning is a functioning of the brain or central nervous system. Where could mental processes be supposed to take place, if they are not bodily processes? And what could the stable mechanisms of sensation, memory, intelligence and desire be, if not physiological mechanisms?

Functionalism, however, is more than the claim that the inner life can in principle be viewed and understood externally or objectively, in the terms of physics and chemistry. It is that the other, distinctive way in which thought and experience can be assessed is not specifically phenomenological, or from the standpoint of the subject, but functional, in terms of a certain pattern of causal relations to other processes and events. The thesis is a radical thesis about intentionality, to the effect that all intentionality is extrinsic. The content of, for example, a belief that there is a sandwich in front of me is explained (to put it very crudely) in terms of the normal cause of a state of that type (i.e. there *being* a sandwich in front of me) together with the normal (when uninhibited) consequences for action of a state of that type (i.e. if I were hungry, other things being equal, I would go to pick up the sandwich to eat it). What it is for the belief-state to have this content is simply for it to be of a type having these causal relations to sensory input and behavioural output.

One reason why this illustration of the view is very crude is the complexity of psychological structure. Notoriously, action is not a simple product of belief or desire, but of belief and desire in combination. Given odd enough desires, almost any behaviour can become appropriate to almost any set of beliefs. Given odd enough beliefs, the same goes for the relationship between specific behaviour and specific desires. Moreover, even if desire is supposed constant, the same belief can produce a wide variety of actions as it is combined in different people with a wide variety of further beliefs; and on the other hand widely different combinations of beliefs can give rise to the same behaviour. So too with desires,

if belief is in turn supposed constant. Finally, there are other determinants of action than belief and desire: capacity and mood, for example. Mere complexity, however, is no obstacle to functionalism, which thrives on reminders that a computer-driven robot designed to replicate human life would have to be internally very complex indeed.

What, then, (to speak to the topic of perception) does the functionalist suppose that we are doing when we give an account of the intentional content of our sensations? What does the subject tell the psychologist studying sensory illusion, or the patient tell the optician, when they describe the way things appear to their senses? The functionalist might approach this question by asking what purpose such talk could serve in the language of a robot constructed to act in all ways like a human being. What significance, with respect to a robot's inner state, could be attributed to its 'report' on a hot day that there appear to it to be pools of water rapidly coming and going on the road ahead?

It seems that any answer compatible with functionalism must relate the 'report' of the robot to a certain kind of inclination to 'believe' that there are pools of water on the road, an inclination founded on a physical state of a sort typically caused by a 'visual' mechanism sensitive to light and operating more or less as it would if stimulated by actual pools of water. The hypothetical 'belief' itself must be explained as the physical basis of a disposition, perhaps inhibited, to act (in accordance with relevant robot-goals or 'desires') as if there were pools on the road. Now the problem of the intentionality of sensation is, to put it crudely, the problem of why such an expression as 'pools of water on the road' should ever enter into a true description of the state of a human being or animal: a problem which might seem especially urgent in that the description can be true even when no pools exist with which the subject might be in relation. The answer to the analogous question about a *robot's* 'sensory state' must be something roughly like this: it is because the 'sensory state' is a physical state of a sort which typically, even if not always, forms an early part of a certain sort of causal process falling largely within the robot, but linking the presence of actual pools of water on the road with robot-behaviour of a kind which is appropriate to the robot's 'goals' only if there are pools of water on the road. In addition (on the hypothesis that it successfully mimics human behaviour) the robot has been so equipped with a self-scanning device that, whenever such a state occurs in it, it is disposed to say something equivalent to 'It looks as if there are pools on the road!' This speech may be otherwise inhibited but, if uttered, can be taken as a 'report' on its own state which is 'true' even if there are no pools, and even if those later stages of the whole process which constitute 'belief' are inhibited.

To the functionalist's claim that this explanation of the significance of a robot's 'sensation report' will equally serve to explain the corresponding human speech, it is natural to respond that for a person such a speech is a genuine report or self-description directly founded on the speaker's state of consciousness. In the language of the argument of foregoing sections, it is the expression of primary and perspicuous knowledge and it achieves primary and perspicuous

reference. It is not a merely contingent accompaniment to a physical process of which the speaker is at best indirectly aware. A functionalist, however, will presumably have passed beyond the reach of such appeals. The strength of his position lies in this: if we accept that events in consciousness are the subjective aspect, so to speak, of certain physical events in the brain, then we must accept that the functionalist's relational characterization of human sensation is at least true, or can be modified so as to be so without theoretical difficulty. The premise that thought is a function of the brain, that consciousness must be physically based, seems now a matter virtually beyond debate. In that case the issue with the functionalist must be, not so much whether he has truly characterized sensation, as whether he has sufficiently characterized it to have explained its intentionality.

Here the argument of earlier sections of the present work constitutes, it may be submitted, an obstacle to functionalism which it cannot easily surmount. A firm distinction was recognized between having a hunch and having a belief grounded on sensation, a belief whose provenance is therefore perspicuous to us. This difference, it seemed clear, cannot be explained merely in terms of the causation of the beliefs, since what is subjectively a groundless hunch might arise as a result of a mechanism of the same general kind as the mechanism of sensation: indeed, in the case of 'blindsight' the mechanism is that of sight itself. The distinction is intelligible to us just because in primary sense-perception the world is presented to us in a unique if familiar way—e.g. in visual sensation—while in subliminal perception, blindsight, extra-sensory perception, the operation of the postulated 'sense of direction' and the like, beliefs or inclinations to believe are produced which are subjectively ungrounded. Such inclinations appear subjectively mysterious or 'intuitive' in the modern sense. They strike the subject as groundless, as indeed they are, even if empirically reliable. All such commonsense explanation of the distinction has to be jettisoned, however, if we are thought to be no different from a robot.

That is not, of course, to deny that a robot could in principle be constructed so as to speak in just the same way as a human being, and so to *seem* to speak of sensations. It could be made to register 'surprise' at some of its abilities to get things right, while offering a complacent explanation of others in terms of 'sensation': 'I know because I saw it.' The robot might even be made to seem to draw the distinction in more explicit and technical terms: 'These beliefs are perspicuously grounded on sensation, those inclinations to believe are just mysteriously reliable intuitions.' A question is, however, whether a coherent and intelligible account can be given of the 'meaning' of 'reports of sensation' in a robot's 'language'.

Let us suppose for the sake of argument that a certain sort of subliminal or extra-sensory perception typically gives rise to confident true belief, and that a robot has been so programmed as to mirror human behaviour in this as in other respects. Comparably to a human being the robot issues 'reports' of the general form (where *s* is a state of affairs): 'I do not see *s*, or otherwise have sensations

of *s*, but I am intuitively sure that *s* holds.’ If we accept the functional definition of what ‘sensation’ *is* for a robot, then it seems clear that the first part of such a report would typically be *false* for the robot, as it would not be for a human being reporting on the confidence provoked by extra-sensory perception. For a state functionally or causally related in the right way both to an object of ‘perception’ and to consequent ‘belief’ would have taken place, *ex hypothesi*, both in the robot and in the human being. That the robot replies as it does, falsely, must be either because its self-scanning system has failed to record the existence of the relevant functional state, or because a true report has been otherwise inhibited. Yet if it did *not* reply falsely, its verbal behaviour would not mirror that of a human being who in such a case would have uttered the same form of words *truly*. The human denial of sensation would have been true because, although the functional state would *ex hypothesi* have existed, it would not have been a sensation. In other words, the purely functional characterization of human sensation is not sufficient because the occurrence of extra-sensory perception, or the operation of any ‘sense without sensation’, will necessarily involve a state which satisfies that functional characterization, but which is not sensation.

The functionalist can of course fight back. He might, for example, deny the possibility of any real distinction between extra-sensory and normal (or, as it is sometimes called, ‘conventional’) perception, simply noting that there are, or might be, cases of perception in which people regularly give sincere but *false* denials that they have sensations. A less heroic line of argument, however, would be to make the functional characterization of robot-sensation more complex, so as to distinguish between ‘unconscious’ and ‘conscious’ functional states of the robot. The latter could be said to be those which are picked up by the robot’s self-scanning mechanism and are therefore such that the robot is disposed to report them. It might then seem possible to argue that ‘sensation’ and related terms in ordinary speech can be explained as denoting, in this sense, ‘conscious’ functional states of the relevant type.⁴⁶

Now the attempt to account for consciousness in terms of a disposition to utter certain sentences might not be very enticing in itself, although it has a not undistinguished history and probably a fair number of adherents today. Such intuitive plausibility as it possesses stems largely from the thought that someone is fully ‘self-conscious’, or ‘conscious’ of their state of mind, only if they have reflectively grasped or appreciated its character: as if anger which is not explicitly recognized as such, since not consciously ‘anger’, is therefore not conscious anger. But if this conception is offered in explanation of consciousness in general, and so of what it is for either anger or a sensation to be (or be manifested in) a state of consciousness, then the surely odd conclusion follows that the distinction between normal and extra-sensory or subliminal perception cannot in principle exist for animals and infants. It follows indeed that living creatures without language cannot enjoy ‘sensation’ in the ordinary meaning of the word at all—although robots with language can. And that would be a paradox

which, like many modern philosophical paradoxes, would seem to assign rather too much significance to the possession of language.

There is a deeper problem. If we try to distinguish in this general way between the perceptual state as a stage in the process of perception (call this state ' p ') and a state of 'consciousness' of p which is a stage in another process giving rise to a report (call the second state ' s '), then it evidently becomes possible, not only that the former should exist without the latter, but that the latter should exist without the former. What, then, is the report itself about? In other words, which state is the sensation? If (as I have imagined being proposed) p is called a 'sensation' in virtue of the occurrence of s , then it is still the case that sensations are states of a kind which might have been unconscious, and which occur in blindsight and the like. If we prefer to link sensations to consciousness essentially, by identifying the sensation with s rather than p , then it becomes possible to have a sensation without being in any state which is typically a stage in a perceptual process. For someone to have an impression of a red cube would simply be for there to be an occurrence in him disposing him to utter a report to the effect there is an occurrence in him which disposes him to believe (roughly, act as if) there is a red cube before his eyes. This mouthful is not an easy one to swallow. First, the problem of accounting for consciousness arises all over again, just because there is a difference, like the difference between primary and extra-sensory perception, between being inclined to utter a report knowing why, and being inclined to utter a quasi-report without knowing why. Second, in cases of normal perception conscious sensation is *as such* a causally efficacious antecedent of perceptual belief. That is hardly surprising, since sensation supplies the perspicuous grounding of the belief. In the case of blindsight, the sense of direction and the like, what is surprising is that a process apparently occurs which is not conscious (although presumably closely related physiologically to conscious sensation), but which gives rise to belief, or to something approaching belief. What is not surprising is that this process is only weakly efficacious: subjects are typically doubtful about their ability to get things right even to the extent that they do.

It seems, then, that there must be something radically wrong with the attempt to deal with the consciousness of sensation by the model of a self-scanner giving rise to reports on the perceptual process. Such a model necessarily fails to do justice to that integrity of conscious experience and thought with our awareness of it which was so frequently remarked in the seventeenth century. As Locke put the commonplace, 'thinking' is an 'Action...that cannot be done without a reflex Act of Perception accompanying it'. A person having a visual impression of a golden sphere is typically both inclined to believe that a golden sphere is present and also to some extent inclined in appropriate circumstances to say such things as that they have a visual impression of a golden sphere. But *both* these inclinations typically arise *just because sensations are conscious states with intentional content*. The first inclination arises in so far as we are normally inclined to trust our senses, and the second in so far as we may have a certain, fairly weak general inclination to give an account of whatever falls within our

cognizance: in this case, an account of how things are presented to us in sensation. When we try to reduce sensory consciousness to those inclinations or their physical bases or antecedents, paradox arises. For sensation is their epistemological, and not merely their material source and foundation.⁴⁷

The possibility noted earlier, that the same physical input through the senses together with the same physical output in behaviour might in principle embrace indefinitely many different internal scenarios, or permutations of belief, desire and the rest, is sometimes supposed to raise a problem for functionalism. How could we ever suppose that we have identified the one scenario which is instantiated in any particular case? But a more profound problem concerns psychological structure in itself. If the mind is in effect just a physical black box, a first difficulty is to see how even the most simplified hypothesis about the structure of its contents should have occurred to us, presumably at a time when we lived in caves and embodied the hypothesis in language. Of course, if we were now to come across a robot acting in a 'purposive' way and 'sensitive' to its surroundings, then provided that neither 'purpose' nor 'information' were too recondite it would be unsurprising if we hypothesized two such specific internal mechanisms acting in combination. Yet that would be because we already see the machine by analogy with ourselves and interpret its output as 'behaviour' analogous to purposive human action suited to the known circumstances.⁴⁸ We have already achieved mastery of the categories *sensitivity*, *belief*, *knowledge*, *desire*, *purpose*, *action* and the like. Now it is plausible to hold that we have these categories, not because of some prehistoric 'folk' speculations about the interior of the black box, but because we ourselves are human and so after all have some more direct knowledge of what it is like to perceive, to believe, to desire, to deliberate and to act.

One might suppose that the only way in which this last thought could be pursued would be through some concession to the traditional assumption that we have immediate knowledge of the contents of our own consciousness. Yet there is a modern approach which avoids this difficulty for functionalism while insisting that the traditional view is a wildly metaphysical distortion of the truth. In fact, so it is suggested, we acquire the categories in question, we learn what it is to believe, perceive and the like, when we acquire language. In the context in which we learn to communicate linguistically, the ascription of beliefs, desires, intentions, intentional actions and so forth goes hand in hand with (if it does not follow after) the intimately related performances of making assertions, expressing wishes, issuing orders and demands, justifying intentions for the future, explaining one's actions by giving reasons for them, giving advice, discussing what to do and so forth. Without a command of the latter sort of performance, it is suggested, a command of the former would be impossible. Thus our possession of the categories of 'psychological structure' is ascribed to our being human ourselves, but to our being social language-users rather than to our having privileged access to our own inner states. In interpreting other human beings or animals in terms of these categories we are trying as it were to capture

them in a conceptual net having its origin in the structure of social communication itself. As language-users we interpret others as, or as if they were, language-users. On this view, roughly speaking, such elements in psychological structure as belief and purpose combine in a seamless whole in our explanations of behaviour just because simple assertion, the statement of intention or desire and the giving of reasons necessarily play interdependent roles in the life of one who speaks a language. It can be argued that only on the assumption of such connections could any system of sounds be interpreted as a language at all. For in trying to interpret the language of a newly discovered tribe we should need to seek out and build on utterances (assertions, wishes, evaluations and so forth) which could be understood as expressions of such beliefs and desires as together make sense of its members' behaviour as rational activity.

One way in which such a 'linguistic' theory of intentionality differs from behaviourism can be put like this. For the behaviourist, actions are prior to words: verbal behaviour owes its meaning to its relationship to non-verbal behaviour. Communication is a type of secondary behaviour which at most complicates the standard account of psychological attributes in terms of stimulus and response. For the linguistic theory, however, possession of a language endows talk of 'behaviour' with a new dimension: as a dog cannot, except in an analogical sense, want to go for a walk with its master, so it cannot, except analogically, go for a walk with its master. But if determinate intentions can be possessed in the primary sense only by those with the means to express them, and if 'behaviour' in the relevant sense presupposes intention, then it seems that linguistic meaning cannot be non-circularly explained in terms of behaviour. There is no way out of the charmed circle of intentional notions. They thus form a conceptual network which is quite disparate from that used by scientists and laymen in giving an account of the natural world and in explaining or predicting natural events.

A difference of opinion among holders of this general approach has, however, emerged with time. Wittgenstein seems to have interpreted the supposed incommensurability of the mental and the physical in such a way that those relations between elements in psychological structure (e.g. between wants and actions) which it is natural to regard as causal are not really causal at all, but explanatory in a different way. A number of his followers, perhaps most eminently Elizabeth Anscombe, have pursued this rather unpromising line of thought, partly by appealing to a traditional distinction between reasons and causes. More recently, however, another advocate of the thesis of conceptual disparity, Donald Davidson, has nevertheless argued that the explanation of action by psychological antecedents or reasons is essentially causal. The incommensurability of the mental and the physical is taken to stand in the way only of psycho-physical laws. The general concept of causality, it is held, operates on both sides of the line, and across it.⁴⁹

Both general approaches to intentionality can be characterized as attempts to find a primary form of intentionality by reference to which the nature of intentionality in general will become perspicuous. The behaviourist or behaviouristic functionalist picks on the intentionality of purposive action, which he regards as untroublesome because he thinks of external behaviour as unproblematically physical. After all, machines can be made which adapt their behaviour to circumstances in accordance with goals.⁵⁰ The linguistic theory, on the other hand, selects the meaning of sentences as primary, taking the representative power of the institution of language to be at any rate less problematic than that attributed to consciousness or the soul. Both theories can appear to have achieved clear progress, whatever difficulties remain for them, simply in bringing intentionality out of the realm of the subjective.

Wittgenstein's approach to consciousness is equivocal, and he sometimes seems to be trying to gain the advantages of abolishing the subjective point of view while not actually doing so. The black box may or may not contain 'consciousness', but in any case, he insinuates, psychological statements cannot be construed as accounts of it. What, if anything, happens in 'consciousness' is irrelevant to the satisfaction of psychological predicates, and it is of the nature of language, which is essentially public, that it should be so. One strand of Wittgenstein's argument can be roughly summarized as follows. We might think (with Locke) that a word like 'blue' has the function of capturing the quality of an experience, just what it is like to have a particular kind of sensation. Yet 'blue' is a word in a public language and its use is founded on common agreement in discriminating objects, not on some privately apprehended quality of subjective experience which, for all it matters, might differ from individual to individual. The argument appears to be directed against the notion of independently identifiable, namable *qualia*; but it implicitly embodies the suggestion that the intentional form of the expression 'sensation of...' (as in 'sensation of blue') reflects, not the intrinsic intentionality of the sensation for which it is employed, but the role of public language (words like 'blue') in, so to speak, signalling sensation. The whole argument famously embodies the claim that an autonomous language for private objects is impossible, and it seems intended to imply that, when we talk of the content of our sensations, we are simply passing over whatever may be intrinsic to them and are somehow relating them to the public world as that world is the object of public communication. For it is in relation to that world, in the full context of social intercourse and common human purposes, that language arises. And it is in relation to that world that such a word in the common vocabulary as 'blue' must acquire its primary meaning. The primary use of 'blue' is for things as socially discriminated, not for sensations; and by no conceivable innovation could the order be reversed for an artificial language.

Some of the strengths and weaknesses of the 'private language argument' will be considered in the section following, with particular reference to Locke's own position. The present point relates only to the wider suggestion that the categories of psychological structure (including the category of sensation) are

reflections of linguistic structure. It is not entirely implausible that the psychological difference and relation between belief and intention is an internalization of the linguistic difference and relation between statements of fact and statements of intention: between such sentences as 'The sun will rise at six o'clock' and 'I will rise at six o'clock'. These 'expressions' of belief and intention do not seem at all like descriptions of one's own state. Yet it seems a short step from 'She will go' to 'I believe that she will go', or from 'I will go' to 'I intend to go'. Even 'I want to go' and 'I desire you to go' do not in general feel like self-descriptions. And yet it is on the basis of such uttered sentences, among other things, that we ascribe determinate beliefs, intentions, wishes and desires to others.

There is, admittedly, a well-known difficulty for this reductive argument which might be called 'the problem of sincerity'. To express a belief or desire is not necessarily to have it, so that the truth conditions of the ascription of a belief or desire seem rather to consist in what lies behind the expression, in the mind, than in the expression itself. The concept of sincerity seems to require the possibility of a fit, or failure of fit, between the intentional content of the mental state and that of its expression. With sensation the difficulty is even more striking. Intuitively, to give an account of a pain to a doctor, or an account of an after-image or bodily sensation to a perceptual psychologist, or an account of what one 'sees' on an optical chart to an optician, is not simply to 'express' a belief or attitude (whatever it might be to do that), but is rather straightforwardly to describe or indicate one's own subjective condition, truly or falsely.

Among attempts to disturb this natural thought there is the notorious suggestion, envisaged (if, perhaps, abandoned) by Wittgenstein, that reports of pain can be regarded as socialized pain-behaviour, articulate groaning so to speak, with the report of the pain's location standing in for less conventional pointings and clutchings. Yet it seems that the latter relationship exists only in so far as pointing (*qua* signal) and even, in its way, clutching (*qua* purposive action) themselves have a sort of intentional content *deriving* (if they are sincere) from the content of the painful sensation. One indicates and clutches a particular tooth just *because* that tooth is the object of the ache. To attempt to explain the tooth's being the object of pain as its being the object of clutching is to put the cart before the horse. Undeliberate groaning, by contrast, has no content, but is a natural response which, although provoked by the quality of the ache, is likely to be independent of the ache's intentional content (although of course a groan might be provoked by the thought that one's leg is broken, rather than by the pain).

Another attempt to represent sensations as linguistically-based constructs, as contents of the black box simply postulated to fit outward-looking utterances, might lay emphasis on such impersonal forms of their description as 'The plate looks elliptical' and 'It seems to be raining'. These might then be construed as qualified versions of 'The plate is elliptical' and 'It is raining', analogous, perhaps, to probability statements. Even the patient's report to the optician

generally takes the form of an account of what seems to be the case out there, rather than an account of the visual impressions discerned within himself. The optician might therefore seem analogous to someone who tests opinion by striking up conversations on the topics of the day rather than by requesting those he meets to introspect their beliefs. The unnatural character of the second technique might be taken to indicate the logical primacy of the first.

One difficulty with any such approach to sensation, however, is that the expressions 'it looks as if', 'it appears that', 'it seems to be the case that' and so forth can indeed perform a general role in qualified assertion (as in 'It looks as if there will be a new government'), but that role may have nothing to do with sensation. To distinguish the use of such sentences in general from their specific and, as we naturally suppose, literal use to describe or report sensations (how things seem to the senses) is more or less to advance in another form the distinction between belief in general and perceptual belief in particular. For their literal use, if truthful, occurs when they are prompted by sensation, and when they derive their content from the content of the sensations which prompt them. We therefore cannot expect to explain the intentionality of sensation by their means. Indeed the possibility of both personal and impersonal modes of the description of sensations, the possibility of both 'I have a visual sensation of a red cube' and 'There seems to be a red cube', simply reflects the essentially representative character of sensation. For it broadly corresponds to the ambivalence or ambiguity ascribed by Descartes to the term 'idea'. There is therefore nothing whatsoever in the role of sensation-reports to lend comfort to the linguistic theory of psychological structure: quite the reverse.

Nevertheless there is a cogent point embodied in the linguistic theory which has a certain indirect relation to Locke's own thinking on two subjects, 'particles' and 'ideas of reflection'. These topics had in effect been linked in a medieval theory which had some influence on seventeenth-century logic. John Buridan had approached the role of syncategorematic words ('is', 'not', 'every', 'or' and so forth) roughly as follows. A categorematic term signifies a thing of which there is a concept in the mind, whereas a syncategorematic word does not signify a thing but expresses an act or mode of conceiving which modifies the relation between the categorematic concept and its object. Thus 'every', in the expression 'every man' expresses an act which modifies the relation between the categorematic concept *man* and men *universally*. In a universal sentence the universal mode of conceiving, as a present act of the mind (*actus exercitus*), is in a sense signified, but '*per modum affectus*'. It is not, that is to say, signified as an object of thought. Yet we do not only have and express universal thoughts, we can think and talk about them in a second-order way. When we do so, what is signified is not the '*actus exercitus*' but an '*actus conceptus*', signified '*per modum conceptus*'. Corresponding to the syncategorematic word 'every', for example, is the categorematic term (and concept) 'distribution'. Similarly, 'affirmation' corresponds to 'is', 'negation' to 'not', 'disjunction' to 'or', 'limitation' to 'only', and so forth. The account was extended to a less logical

subject-matter, to cover the relationship between the expression of joy, 'hurrah', and the name; 'joy'; or between the expression, 'if only...' and the name, 'desire'.⁵¹

All this has echoes in the *Essay*. In his theory of particles, as we have seen, Locke adopted the standard view that they are 'marks of some Action, or Intimation of the Mind' in prepositional thought and in reasoning. With respect to the latter, they '*shew what Connexion, Restriction, Distinction, Opposition, Emphasis etc. [the speaker] gives to each respective part of his Discourse.*'⁵² In arguing that 'but' is multiply ambiguous, Locke finds that it expresses sometimes a certain 'stop of the Mind', sometimes a sort of 'limitation', sometimes an 'opposition', sometimes a form of 'conjunction' and sometimes, apparently, actions or 'intimations' which lack names altogether.⁵³ All these names must be supposed to stand for 'ideas of reflection', i.e. for ideas of the operations expressed by the particles. We can thus see how Locke could hold that ideas of reflection come to us late, and only if we pay direct attention to those operations which even a child can express. The child who properly employs the conjunctive 'and' does not thereby have the concept of conjunction, any more than the child who joyfully shouts 'hurrah' thereby has the concept of joy.

The Buridianian theory is by no means crude, but it nicely illustrates the contribution of language to conceptions of what goes on in the mind. The classification of constructions according to their logical point is taken to be a classification of mental operations on concepts. Yet once we see that such operations as 'distribution' or 'limitation' owe their very possibility to language, the need to think of them as species of mental act seems to dissipate, and the move to the psychological level to lose any explanatory force. Nothing is gained by attributing our ability to employ the word 'not' to an ability to indulge in the mental activity of negating: the concept of negation is the concept of a linguistic operation, not of a non-verbal process in consciousness. The point can be extended plausibly enough to non-logical contexts. Sincere cursing or sincere praying is not to be explained as the expression of a peculiar process in consciousness, the real or essential cursing or praying, which might have occurred without its expression. Cursing and praying are essentially verbal activities, and they are sincere, roughly speaking, when they are performed in a certain context of belief and action by someone who grasps their significance. A similar point applies to at least some of the distinctions we draw among motives or emotions. The expression of scorn is essential to scorn, and a being without language lacks the means to be scornful.

What I have called the linguistic theory of psychological structure is in effect the bold claim that such an argument can be pressed to the hilt: if we abstract from language, *all* psychological structure withers away, and, without structure, 'the mind' withers away. The mind is just an inwardly cast shadow of the external employment of language, and the traditional enterprise of considering states of consciousness or 'modes of thinking' in abstraction from language is entirely misconceived. So at least it might be argued, given a certain amount of

nerve. Yet the argument promises more than it can deliver. Whatever the power of language to mould thought and experience, and however much the structure of discourse determines the ways in which we slice up the mind and its operations, there is something prior to language which makes language possible. Unless there were animal experience and intelligence and emotional life prior to language, there would be nothing to incorporate and comprehend language, or to be transformed by it. Our power to use language cannot, therefore, be the explanation of the intentionality of our mental states.

Private language and secondary qualities

Although a certain view of sensations and sensation-reports which seems ascribable to Wittgenstein was rejected in the foregoing section, it does not follow that there is nothing to be learnt about them, or about the shortcomings of Locke's conception of them, from the Wittgensteinian argument against the possibility of a private language. At the same time Locke's view has certain strengths by comparison with Wittgenstein's. The two may perhaps be brought into productive conflict by the consideration of a well-known, but at first sight rather surprising passage in the *Essay*.

When in the grip of his conception of ideas as blank sensory effects, Locke proposed, as we have seen, that such words as 'blue' are ambiguous, properly or primarily standing for the effects and only by a misleading, if not illicit extension for whatever aspects of objects cause those effects. Fire, he intimated, is improperly 'denominated...Light and Hot, as if Light and Heat, were really something in the Fire, more than a power to excite those *Ideas* in us'.⁵⁴ He did not always retain this sense of an impropriety in the use of the names of ideas of secondary qualities for attributes of objects themselves. When stressing the truth of all simple ideas, he wrote that 'the Name *Blue* notes properly nothing, but that Mark of Distinction, that is in a *Violet*, discernible only by our Eyes, whatever it consists in.'⁵⁵ But it seems central to his official view that 'blue' has two employments: in the first instance it is the name or sign of an idea, while secondarily and derivatively it names whatever in the object causes the idea, and is therefore signified by it.

The argument for the unimpeachable 'truth' of simple ideas, however, was developed at one point in what might seem a contrary direction. Locke considered the possibility that 'by the different Structure of our Organs' different ideas of colour should be produced in different minds by the same intrinsic attributes. He noted that, if this difference between people were wholly systematic or symmetrical (as it is not, of course, in ordinary cases of colour-blindness), everyone might still make the same distinctions between things as other people. Their different simple ideas would signify the same intrinsic attributes, marked by the same names: if one person's ideas of blue and yellow were the reverse of normal, 'neither the Ideas hereby, nor the Names, would be

at all confounded, or any *Falshood* be in either'. The net effect on his practical application of the words 'blue' and 'yellow' to objects, or on his capacity for understanding, would be zero.⁵⁶

Now this argument might suggest to the modern reader a very different view of the relationship between the supposed two uses of colour predicates, for ideas and for objects. No longer is it so clear that Locke saw the latter use as dependent on the former. For in Locke's story two men know qualitatively different ideas as 'the idea of blue', so that, in its application to ideas, they mean different things by the name 'blue'. If the use of 'blue' for objects were secondary and derivative, and meant for each man something like *having the power to cause the idea of blue*, then it would follow that 'blue' meant different things for each of them in this secondary use too. Yet it may seem to be the burden of Locke's argument that, as far as the public use of colour predicates for objects is concerned, what matters is that all people distinguish objects according to their intrinsic attributes by the visual effects of those attributes, whatever those effects might be in each person. Hence the public use seems to be cut free from the private one, the former depending only on agreement in judgement as expressed in a shared language. The phenomenal quality of the experience upon which the judgement of each person is grounded becomes irrelevant.

Such a line of thought, if it could be ascribed to Locke, would have a particular interest because of its seeming connections with a strand of the argument against the possibility of a private language. On that argument, the two uses of 'blue' are indeed systematically connected, but its primary use is for publicly observable objects, surfaces, light and so forth, while it is its use in characterizing visual sensations which is secondary. 'It looks (seems) blue' has, it is reasonably argued, a meaning parasitic upon 'It is blue'. For, if we are to describe the 'phenomenal quality' of experience, as doubtless we can do, we must use words from the public language which we all speak, words which derive their primary sense from their application to publicly observable objects. 'Blue' cannot be supposed to have a different or underived 'private' sense as used for sensations. It follows, so the argument continues, that a story like Locke's is actually incoherent. For the people whose colour sensations are (according to his hypothesis) symmetrically transposed will in general agree with other people in their judgements not only over which objects *are* blue or yellow, but over which objects *look* or *seem* blue or yellow in particular circumstances, whether normal or abnormal. The same goes even for their judgements of the form 'It looks blue, or seems yellow, to me'. With such community of judgement, so the Wittgensteinian argument continues, goes community of meaning. Consequently we must admit that after all violets and marigolds give rise to the same visual sensations, namely sensations of blue and yellow respectively, in both purportedly differentiated sets of observers. No difference in the phenomenal quality of the experience of such observers can be described or indicated in language. Contrary to hypothesis, therefore, no such difference exists. Locke's story is incoherent.⁵⁷

Whether or not this reasoning is sophistical (as no doubt it is), it would be a mistake to suppose either that Locke himself came in the least near to it, or that he proposed conflicting models of how predicates like 'blue' possess meaning. For what I have said that his argument might suggest to a modern philosophical reader is neither explicit nor implicit in it. He neither recognized anything but a tight link between the supposed two uses of 'blue', for ideas and for objects, nor went so far as to conclude that his two people with systematically different ideas apply 'blue' to external objects with the same meaning. His conclusion was only that the old sceptical speculation that people's colour sensations perhaps differ systematically, even 'if it could be proved', would make no difference 'for the Improvement of our Knowledge, or Conveniency of Life'. He was not after all inclined to argue that the meaning of 'blue' in its use for objects is 'public' in some strong sense, or to deny that it is tied for each person to the phenomenal quality of a particular way of experiencing things. Still less was he suggesting that the use of 'blue' for objects is primary, and its use for ideas secondary and derivative. He was, rather, attempting to make the same, familiar general point as before, so central to his epistemology, that all that matters in the idea of a secondary quality, as far as its role as sign and representative (or its 'truth') is concerned, is that it has *some* distinguishing character. As he says of such ideas in this particular argument: 'each of them being in the Mind, such as it is, suitable to the Power that produced it, and which alone it represents, it cannot upon that Account, or as referr'd to such a Pattern, be *false*'.⁵⁸ Having drawn that familiar moral out of the argument, he was careful to make it clear that he did (very reasonably) believe that people's senses in fact function in similar ways to similar effect. It is for that reason, no doubt, that he believed that different people do after all mean the same thing when they call objects 'blue'.

There is in its context nothing particularly surprising about Locke's belief that the primary use of secondary-quality predicates is for sensations rather than things. His explanation of the ontological status of secondary qualities as *powers* seems to presuppose that the identification of the defining effects of these powers is prior to the identification of the powers themselves. To reject his notion of simple ideas as blank effects is implicitly to deny that presupposition. As it has been argued above, to characterize a sensation even as a sensation of red is to characterize it in terms of an aspect of intentional content: i.e. in terms of a way in which *things in space* appear or are presented to us. More generally, to give the content of a sensation is to give an account of the conditions under which it would be veridical. Expressions like 'a feeling of weight', 'a visual impression of a circle', 'a sensation of blue', 'the appearance (to me) of a speck in the sky' embody such accounts. Words like 'weight', 'circle', 'blue', 'speck' enter into these expressions, not because there is a sensation or sense-datum in question which is heavy or circular or blue or a speck, but because, if the sensation is to be veridical, then something physical must be causing it which is, respectively, heavy or circular or blue or (at least from here and without a telescope) a speck. That point may not show that the characterization of

secondary qualities as powers is false, but it does show that such a power must be an epistemologically very peculiar power. So much, perhaps, is only to be expected of a power which is directly and primarily manifested in a feature of the appearance of its possessor to observers.

The 'private language' argument is, then, entirely right in maintaining that predicates like 'blue' are primarily predicated of objects rather than sensations. Yet it is objectionable in so far as it maintains that the sameness of the phenomenal quality of different people's experience is secondary to the sameness of what they would say about it: to community of judgement. Whatever its weaknesses, Locke's account evinces a firmer grasp than Wittgenstein's on what we might call the ontology of perception. In particular, it allows proper significance to the necessity for a sense-mechanism, a physical embodiment of the faculty and of the 'regular ways' in which objects affect us through the sense in question. It is that necessity which carries with it the possibility that, with a different mechanism, the sensory effect of the same intrinsic attributes might be different. There is nothing inherently or irremediably sceptical about Locke's speculation of transposed colour-sensations. It is just that he took himself to be ignorant of the relevant sense-mechanisms. Knowledge of those mechanisms, he reasonably assumed, would in all probability falsify such speculations rather than confirm them.

An illustration of this specific virtue of Locke's approach by comparison with Wittgenstein's is provided by something more like ordinary cases of colour-blindness. For the Wittgensteinian view, at least on a certain popular interpretation of it, to have the concept *blue* is to have a discriminatory capacity: people who make the verbal discriminations which 'blue' and its synonyms mark in public language have the concept. Given that capacity, there cannot be any reason to deny that blue things appear blue to them. For any further question about the quality of people's subjective experience beyond what can be thus settled at the level of inter-subjective judgement is held to be meaningless. One problem for this argument is that of explaining why the relevant common judgement of colour needs to be stimulated by vision for all who participate in it. Passing over that difficulty, however, we can turn to another. If some people could make all other colour-distinctions, but could not distinguish between red and green, then as far as red and green are concerned they could reach non-accidental agreement with others on the immediate basis of vision only at the level of such judgements as 'This object is red or green'. It would seem to follow from the Wittgensteinian position that the question whether such people see red as green or green as red would be meaningless. Yet, on a view of the matter which is both sensible and straightforward, full knowledge of the peculiarities of their sense-mechanisms might even give us reason to distinguish these red-green colour-blind people into two classes, those who see green things as red and those who see red things as green. Speculation about the phenomenal quality of experience which goes beyond mere community of judgement does not seem in the least senseless. Such an account of experience need not be locked in an

impossible, irremediably private language. Indeed a colour predicate, in connoting a phenomenal attribute, is if anything more closely tied to a peculiar mode of visual experience than even Locke suggests, since it is not necessary that the single secondary quality or power should consist in all its instances in a single intrinsic attribute. It is said that quite separate ranges of light produce the same range of colour sensations, and that, for example, the quality brown ‘in objects’ consists in quite different types of electronic structure.

This latest consideration counts against a well-known proposal by Saul Kripke, who suggests that colour-predicates have ‘a certain referential element’ and are to be understood by analogy with predicates connoting membership of a natural kind such as ‘gold’ or ‘horse’, or, indeed, with proper names. The meaning of the noun ‘gold’ (*pace* Locke—the issue will be discussed in [chapter 30](#), below, as well as in Volume II) is not encapsulate in any such definition as ‘stuff which is yellow, heavy and malleable’, although this latter complex predicate could be used in explaining what the meaning of ‘gold’ is. Like indefinitely many other such predicates, it could be employed to identify the type of substance of which ‘gold’ is the name. Kripke puts this point by saying that such a description does not define, but may fix the reference of the term ‘gold’. Similarly, so he suggests, the reference of ‘yellowness’ is fixed by the description ‘that external physical property of the object which we sense by means of the *visual impression of yellowness*’, or ‘that (manifest) property of objects which causes them, under normal circumstances, to be seen as yellow’. That is to say, the reference of ‘yellow’ is for us fixed ostensively through a certain sensory appearance, yet its meaning is not tied to the appearance, but to what in the world normally causes the appearance. So it is accidental that yellow objects as such look as they do, for ‘if we had had different neural structures... if we had been blind and so on, then yellow objects would have done no such thing’.⁵⁹ In such a possible world, given no intrinsic difference in the objects, it would still have been true that the objects were yellow.

For Locke, on the other hand, a possible world without sight or a sense of smell would be a world without colours or stinks, just as a world without locks would be a world without keys, even if for some reason there accidentally existed in that world pieces of metal physically indiscriminable from keys in the actual world. Now on this particular issue Locke’s model is preferable, and one argument which can be used against Kripke is the point that the unity of the class of brown things, or of sweet things, or of acrid things, might very well exist only at the phenomenal level. Such a supposition is not, after all, like supposing that several very distantly related species of animal have hitherto passed as a single species. For that state of affairs, if discovered, would make continued use of a single name loose or ambiguous, rather like the use of ‘stone’ for both diamonds and gall-stones. To put the contrast crudely, ‘yellow’ is about as much like ‘speck’ as it is like ‘carbon’. What essentially unites specks is a certain look, and something similar is true of yellow things. They differ in that there is virtually no physical restriction at all on what physical objects can appear as specks at the

right distance, and no distance is the 'normal' distance from which to view things. Consequently there is not the same room for a distinction between looking a speck and being a speck as there is between looking and being yellow.

Kripke perhaps sees his suggestion as an explanation of how it is that the primary use of 'yellow' is for objects rather than for a sensation. In that case it is unnecessary, since an explanation is already contained in the point that 'the sensation of yellow' is just an aspect of the sensory awareness or representation of things in space. To perceive the quality yellow is not to perceive a pure *quale*, not even, as Locke supposed, to perceive a pure *quale* evidently caused from without. It is to perceive a spatial object *as* yellow. Nevertheless, *within* the framework of the perception of things in space, to see something as yellow rather than as blue just is to be affected by it in one rather than another specific way within the range of a peculiar mode of sensibility. In tying 'yellow' rigidly to *what* impinges on us, Kripke wrongly breaks this semantic link with *how* it impinges on us.

Yet if we turn to other Lockean examples of secondary qualities, Kripke's approach seems more satisfactory. A possible world without sensation is not thereby a world without heat and light, as it is a world without colour and smell. That might seem surprising, given the traditional connection between light and colours as the 'proper objects' of sight. It is, moreover, a common empiricist thought that all these terms should be regarded as ambiguous, each having not only an ordinary pre-scientific meaning linked to sensation but also a scientific meaning bestowed by theory and the possibility of objective measurement. Yet there is no sense of 'heat' in which the existence of heat depends on the existence of consciousness, and no ready sense of 'yellow' (unless in the amalgam, 'yellow light') in which the existence of yellowness is not so dependent. There is nothing odd in the notion of imperceptible heat or light, but there is in the notion of imperceptible colours and smells. 'Ultra-violet' and 'infra-red' are not the names of imperceptible colours but of types of light-heat beyond what is visually perceptible and coloured.⁶⁰

The explanation of the difference between heat and colour seems to lie, as one would expect, at the level of our pre-scientific experience. Heat is felt, but is also what dries off moisture, melts ice, makes pots boil, causes burns and can be transmitted from one object to another. The thermometer was not the first means available for recognizing that objective or physical heat goes beyond, and can in some sense be distinguished from, heat as far as we feel it. In the nature of the case human beings were never without such means. It is about as difficult to accept that 'hot' sometimes connotes, or might easily have connoted, a purely phenomenal or sense-relative attribute as it is to suppose that 'hard' or 'solid' do so.⁶¹ Light is perhaps less observably and variably efficacious than heat, and so is like colour at least in being ordinarily identified more or less exclusively by sight. Yet, like heat, it is experienced as emanating from particular sources: it casts shadows and is sometimes visible in perceptibly discrete rays. It is what

makes it possible to see things, an identifiable physical condition of vision as well as an object of vision. Colours, on the other hand, are ordinarily identifiable simply as aspects of objects of vision. Hence the meaning of 'yellow' is tied to how we perceive things, as the meaning of 'hot' and 'light' are not.

Despite such differences between them, there is still some reason for considering the Lockean secondary qualities together in the context of a discussion of perceptual content. For the possibility of different levels of description, observational and theoretical, itself generates the common problem of what we perceive when we perceive such qualities. If an observer feels the poker's heat, there seems to be a sense in which he feels its molecular motion and a sense in which he does not. So too with yellow: there seems to be a sense in which the electronic structure in question has a characteristic appearance, and a sense in which it has no appearance at all. The ambivalence arises because, although the theoretical description of heat and colour identifies what it is in the object that impinges on the senses, how it strikes them, or how it appears, can only be given by the observational description. It is a natural thought that we perceive colour and heat but do not perceive electronic structure and molecular motion: yet the thought cannot be maintained without qualification, because we do feel molecular motion, namely as heat, and we do see electronic structure, as colour. A partly analogous case is that of seeing a man as a speck and not as a man; but, whereas there is an experience of seeing a man as a man rather than as a speck, nothing counts as feeling molecular motion as molecular motion rather than as heat. Moreover, for the reasons suggested in the previous paragraph, that analogy is rather closer in the case of colour than in the case of heat. There is more point to the paradox that colour (like a man's being a speck) is 'nothing but' a sensory or visual effect than to a parallel claim about heat. The identification of heat with molecular motion seems unconditional or (in the relevant ways) unproblematic, so that heat is something unequivocally in the object. Colour is (say) electronic structure as seen, and electronic structure is (if anything is) colour as it is in the object, but it seems to do less than justice to the situation to identify colour with electronic structure (or whatever) *tout court*. In the terms of Kripke's discussion, there are no possible worlds in which heat is something other than molecular motion, whatever in those worlds might feel as heat does in our world. But there are conceivable worlds in which yellow and red exist, but consist in different structures from those in which they consist in the actual world. For we can imagine that different ranges of light should have given rise to the characteristic appearances of yellow and red.

More revealing with respect to the fundamental flaw in Locke's theoretical framework is, as we have seen, his influential analogy between heat and pain. He asked how it could be thought reasonable to treat heat or light as a quality of the fire rather than simply, like pain, as an effect on us. It is not unusual for modern philosophers to try to answer this question on broadly the same assumptions as those on which Locke asked it. Pain, we are told, is not as regularly associated with types of object or with their other properties as heat is. The power of objects

to cause physical pain depends on extraneous circumstances, such as the way in which they happen to collide with us. Almost anything in almost any condition can cause pain in the right circumstances, and so forth.⁶² These answers look for the difference in the causal conditions of the sensations in question, the sensations themselves being treated as Locke treated them, as blank effects. Yet if we turn to the intentional content of the sensations, then we shall find (as it has already been suggested above) that physical pain is as invariably felt to be 'in' an object as heat is. We *do* think of pain as 'in the object', because that is how we experience it. That object, however, is in each case a part of our body itself, the phenomenal location of the pain. Pain is a mode of awareness of our body, not of objects external to the perceiver, whatever role those objects may have played in the causality of the pain. It is the burnt hand, not the fire, which enters into the intentional content of the sensation of pain. By contrast, not only our body but whatever is in contact with our body (solid objects, surfaces, the air, liquids etc.) can be the intentional objects of feelings of heat and cold. Light always appears 'external', for although there may be experiences describable as appearances of light in the head, they are not sensations 'of' the head, as headaches are. These differences in intentional content lead to other differences, such as the duality within the concept *pain*. The words 'pain', 'ache' and so forth serve not only for what is 'in' the object, the secondary quality of the toe or tooth, so to speak, but also for the sensory state which has that quality for its object, i.e. the peculiar mode of experiencing the toe or tooth. Hence, although a pain can be illusory (as transferred pain, or pain in a foot which no longer exists), we cannot feel ourselves to be in pain without being in pain, as we can feel hot without being hot. *Pace* Locke, neither 'heat' nor 'yellowness' are ever, like 'pain' and 'ache', names of sensations themselves.

If it seems a little daring to describe a pain in the foot (even in the appropriate sense of 'pain') as a secondary quality of the foot, it may be worth considering why. Is it because there is no presumption that a certain single underlying physical condition is present in any part of the body in which pain is present? That this explanation is inadequate as it stands is clear from the possibility that all acrid fumes or even all brown surfaces do not share some one general physical property in virtue of which they smell acrid or look brown; and yet none the less all are acrid or brown. What can be said is that in each particular case of the acrid or the brown there is *some* general basis for its appearance as acrid or brown, even if that basis is not universal to all acrid or brown things. In other words, there are at least *species* of the acrid and the brown. Even if the only universal account possible of the physics of the brown is disjunctive, there is a presumption that the disjunction is not indefinitely long. If we let into the account *any* physical state of an object which happens on occasion to stimulate sensations of brown, then the distinction between being brown and merely looking brown on some occasion, a distinction essential to the conception of brownness as an objective attribute, would be undermined.⁶³ Yet *anything* which stimulates pain-receptors will do as the physical basis of the pain. Transferred

pain and pain in a non-existent limb are illusory not so much because the sense-organ has been stimulated by something inappropriate, as because the receptors in the part of the body presented have not been stimulated or involved at all.

Some philosophers, for example Reid, have suggested in effect that pain does have a general objective basis, namely physical damage. If that were thought to justify regarding pain as a full-blooded secondary quality, then it would have to be held that any pain in a foot which is felt when the foot is undamaged is illusory, as if there were not really a pain in the foot. Yet that is not enough for a pain to be illusory. Any pain consequent on the activity of receptors in the part presented is veridical, not only if they were stimulated by some identifiable or unknown state of that part (whether 'damage' or not), but even if the activity constitutes a malfunction of the receptors themselves. Yet that an 'appearance' which is due to a malfunction of a sense-organ should nevertheless be veridical seems blatantly at odds with its being the appearance of any objective attribute at all. In fact a rather similar issue does arise with respect to colour, in the case of so-called physiological colour. That, to normal observers in normal light, the same surface should regularly look brown in one surrounding context and yellow in another is put down to the physiology of vision. For that reason (although there is no malfunctioning) it is impossible to accept that there is a 'real' difference or change of colour in such a case.

However that may be, it is obvious that purely causal considerations are insufficient to explain why heat is attributed to external objects, but pain is not. So much is demonstrated by the fact that, as Locke no doubt had in mind when he endorsed the analogy, feelings of pain are quite as dependably, even if not as frequently associated with fire as are sensations of heat and light. Mild electric shocks give a further illustration of the same point. They cause extremely characteristic sensations and do so in an absolutely dependable way, yet in receiving a shock we are not feeling or perceiving by touch either the electricity in the wire or the wire itself. We are suffering the effect of the electricity. For the shock involves an acute awareness of a part of the body, not of the wire or of something flowing through it. It is the intrinsic intentional content of the sensation, not its causality, which determines that neither the wire nor the electricity is in itself an object of sensation or perception.

Locke's doctrine as to the primary meaning of secondary-quality predicates may seem, as I have said, entirely consonant with his characterization of secondary qualities as powers. Since powers are identified through their effects, it may seem that the effects must be first and independently distinguished. Yet because the relevant effect, in the case of a quality like yellow, is a way in which objects appear or are presented to us, that effect *cannot* be distinguished independently of the quality in the object. The effect, the visual impression of yellow, has to be distinguished in terms of the phenomenal attribute, rather than the reverse. Nevertheless the characterization of the phenomenal attribute as a power is not simply false or inept. This rather complex situation seems to be interestingly reflected in a certain tension in Locke's own arguments which has already been

examined above. For although by his doctrine of the primary sense of 'yellow' he seems to have held that his characterization of yellowness in the object as a power is a strict semantic analysis of a secondary sense of 'yellow' in terms of its primary sense, his general theory of representation implies that the simple idea of the external quality of yellow is more primitive than the idea of power itself. At any rate, it does not seem that the appeal to the idea of power in the account of secondary qualities was an appeal to something of which an independent account had been, or even, on Locke's principles, could be given.

Part IV

Particulars, universals and intuitive knowledge

Introduction to Part IV

Part IV will be something of a miscellany compared with previous Parts, treating such disparate topics as space and time, meaning, the imagination and necessity. Yet there is a line of argument which connects them. The main purpose is to examine Locke's answer, and to suggest a different answer, to what is perhaps the most puzzling question in epistemology: the question of the nature of *a priori* knowledge and necessary truth. To get that far it will be necessary to consider his account of the distinction between particular and universal thoughts, in comparison with other accounts. Since he took it that the latter are distinguished from the former by their abstracting from 'considerations of time and place', it will be convenient to preface discussion of these matters with an examination of his general account of time and space, and our ideas of them. By doing so we may at least avoid the assumption that thinking about particulars is unproblematic, and that only thinking about universals raises any problem.

Present-day theories about the nature of *a priori* knowledge fall for the most part under one of two heads. On the one side, in accordance with the tradition of logical empiricism, it is still held by some that such knowledge is essentially linguistic, and that necessary truths are 'analytically' true, by definition of the terms involved. On the other side, that general proposal has been widely rejected, largely on the ground that it is impossible to justify a sufficiently tight conception of the meaning of a term to support a firm formal distinction between analytic and synthetic propositions or sentences. The most that can be said in favour of the distinction, it is commonly held, or in explanation of the propensity of philosophers to draw it, is that there are some sentences that we are (and would be in almost all circumstances) particularly loathe to give up as we trim and modify our system of beliefs under the pressure to adopt that system which constitutes the most coherent and useful response to sensory input to date: useful, that is to say, for the purpose of prediction.

Both explanations are supposed to cover three traditional distinctions: between the analytic and the synthetic, the *a priori* and the *a posteriori* and the necessary and the contingent. Neither is particularly plausible. The first, the linguistic theory, does not seem to do justice to the distinction between grammatical, and logical or conceptual constraints. The second suffers from, if anything, even

greater difficulties. First, it does not seem that any satisfactory account of 'coherence' can be given which does not itself presuppose some firm *a priori* relations. Moreover, *within* each category of beliefs, mathematical, biological, physical or everyday, there will doubtless be some beliefs that an individual would more readily give up than others. Indeed, many people would more readily give up the belief that there is no greatest prime number than the belief that most people have ten fingers. That does not affect their recognition that, *if* it is true that there is no greatest prime, that truth is necessary, whereas the proposition that most human beings have ten fingers, whether true or false, is contingent. Similarly, if the former can be known to be true or, for that matter, false, its truth or falsity can be known *a priori*; whereas the truth or falsity of the latter can only be known *a posteriori*, from experience. There may be many disputable questions as to just how these distinctions fall (e.g. the question whether they all three fall at the same place), but it should be clear that they concern types of propositions or belief irrespective of the truth or falsity of particular instances of each type, or the degree to which we, or human beings in general, are committed to their truth.

The approach adopted in the present work therefore bypasses this modern heresy, returning for its starting-point to a traditional dispute, in which Locke himself was engaged, about the role of language in *a priori* reasoning. On the one hand it seems impossible that such reasoning should occur without symbolism, but on the other the rationality involved seems to be more than merely linguistic competence. Rational insights seem different from merely linguistic intuitions. The exploration of this issue will require as its preliminary some reflection on the nature of linguistic meaning and understanding, and on the role of intelligence in language acquisition. Previous conclusions about knowledge, and in particular 'primary' knowledge, will be brought to bear on the task of understanding both linguistic understanding and *a priori* insight. The result will be an account of *a priori* reasoning and knowledge which is like Locke's in certain fundamental respects (even, perhaps, sufficiently like it to deserve the name of 'intuitionism') but which is unlike it in recognizing the indispensable role of language in the possession of such knowledge, and the essential fallibility of human intelligence.

Locke's arguments on space and time in context

There was general agreement in the seventeenth century with Locke's principle that '*Where and when are Questions belonging to all finite Existences*',¹ although Cartesians and some others restricted the principle to *material* particulars. There was less agreement, however, about the ontological status of the containers, space and time, and the disputes about them were of considerable historical and philosophical importance. Locke's thoughts on the question may appear peculiarly rambling and inconsequential unless they are read as, in effect, a selective commentary on earlier writings. It will therefore be helpful to take a brief look at the doctrines of Aristotle, Descartes, Gassendi, Hobbes and the Cambridge Platonist, Henry More. There is also the contentious question of the possible influence of Newton.

A chief motive of Aristotle in his account of space seems to have been a desire to avoid the view that space is a permanent extension independent of bodies and prior to them. No doubt he wanted to distance himself from Plato's model of space as the receptacle of forms, a role fulfilled in his own philosophy by matter. Nevertheless he started with the admission that a body's place is something distinct from it, and so can be neither its form nor its matter. For in movement a thing changes place, leaving its original place behind. He concluded, roughly speaking, that the place of the body is the inner surface of what contains it. But a body may move even if it does not leave what contains it, like water carried in a bucket. Movement, and so place, is therefore always relative to some container (not necessarily the immediate container) regarded as motionless. The water in the bucket moves because the bucket moves in the air. The notion of empty space as a possible container Aristotle rejected as absurd by a series of arguments vitiated, from the point of view of the New Philosophy, by their reliance on the supposed explanatory force of his own theory of natural motion and the proper places of the four elements within the lunar sphere.²

Time is approached in a similarly reductionist spirit, as an abstraction from change and, in particular, from movement. Although the argument is extended to *kinesis* in its broader sense of change in general, Aristotle seems to have thought that the temporal 'before' and 'after' can be explained as an application or extension of the spatial 'before' and 'after'. The notion of time arises when we

apply number to uniform motion, i.e. when we numerically distinguish parts of a uniform motion as 'before' and 'after'. Each 'now' is a boundary between such parts which both unites and divides them as a point unites and divides parts of a line. By such points we can divide motion into units and so measure it. Aristotle allowed that the same measure can be applied to rest, but whatever is in time, whether in motion or at rest, is at least in principle movable and subject to change. Immutable things are not in time, but are eternal. Thus time is defined as the 'number of movement in respect of the before and after'. *Because* time is the measure of uniform movement, uniform movement, such as the movement of the sun, is also the measure of time.³

Descartes arrived by a very different route at a conclusion broadly like Aristotle's, i.e. that space and time are nothing over and above things in space and time, and that the void is impossible. Since for him extension is the essence of body, it follows that space cannot be ontologically distinct from body. True, the extension of a body is logically distinguished from that of the space it temporarily occupies, but that, according to Descartes, is only because we artificially treat the latter as a kind of genus which different particular bodies may successively instantiate by sharing the same size, shape and relation to other bodies. Here he allowed a distinction between 'internal' and 'external' place. In considering the former, i.e. the space which a given body occupies, we are focusing on the body's dimensions. In considering the latter we are focusing rather on its relation to other bodies, so that 'external place' can be defined more or less as Aristotle defined place. The argument against a vacuum is simple. Our idea of space as something extended is the same as our idea of matter. Space is therefore real and substantial. The idea of empty space, or space which is nothing, is, Descartes held, a contradiction. What is commonly called empty space is like the air, matter which lacks sensible accidents. Another of Descartes' claims became particularly notorious: 'If you make the supposition that God removes all the air in a room without putting any other body in its place, you will have to suppose *eo ipso* that the walls of the room will touch each other.' He presumably meant that as the air was extracted the room would necessarily collapse, but his proof that 'two bodies must touch when there is nothing between them' reads like a verbal sophism.⁴

An issue which became insistent for Descartes was the question much debated by the Scholastics as to whether the created world is finite or infinite in extent. In general, space as such had seemed infinite, since, whatever boundary is conceived, it is possible to imagine something beyond it. Yet it also seemed possible, not to say theologically necessary, that the material world should be finite. An orthodox doctrine was that the finite world was created within so-called 'imaginary space', an uncreated nonentity with dimensions conceived by analogy with sensible bodies. For Descartes, however, to imagine space beyond a boundary is to imagine matter there. He concluded that we cannot but conceive matter as boundless, but he refrained from pronouncing it infinite on the ground that it may have bounds beyond human comprehension.⁵

Descartes distinguished time from duration, which he took to be prior to, and independent of, change, although change itself has duration. As space is nothing distinct from matter, duration is nothing distinct from things in general considered in so far as they continue to exist. We measure their duration by the duration of 'the greatest and most regular motions, which are those that create years and days'. The divided duration of such motions we call time.⁶

Hobbes' position, as one would expect, is rather different. Our conception of space is not the same as our conception of body, for 'the same space contains sometimes one, sometimes another body'. Yet like all ideas, it must be acquired from body through the senses. It is in fact the purely abstract conception of bodies as external existents: '*Space* is the phantasm of a thing existing without the mind simply.' Supposing the external world annihilated, the conception of empty space which we could form would derive from our memory of external objects as such: 'an imaginary space indeed, because a mere phantasm'. This idea is interpreted by Hobbes as the idea of the *possibility* of three-dimensional objects: for 'no man calls it space for being already filled, but because it may be filled'. Space is nothing actual outside the mind, but to consider it is to consider the possibility of bodies independent of the mind. Hence we cannot deduce that the world is infinite from the indefinite extension of space.⁷ Moreover the notion of a vacuum within the world makes sense. A vacuum will exist between two bodies 'if another body may be put between them, that is, if there intercede any imagined space which may receive another body'. The *a priori* arguments against a vacuum are trenchantly dismissed. No man 'in his natural senses' could believe that two bodies necessarily touch when there is nothing between them. The argument that a vacuum cannot exist because it is nothing is as 'childish' as an argument that 'no man can fast, because to fast is to eat nothing; but nothing cannot be eaten'. Empty space, although nothing, may need to be computed in computing the relationships between real and independent bodies.⁸

As it happens, however, Hobbes equally rejected the arguments *for* a vacuum within the world, taking it on experimental grounds that there is in fact no space between bodies. The question whether the world is infinite, or whether there exists space beyond body, 'can never be attained by a finite inquirer'.⁹ But even if all space is supposed full, Hobbes seems to see the notion of space as having a fundamental role in an ordered conception of the world. Every body (i.e. every substantial thing) is conceived of as a *subject*, he tells us, because it is '*subjected* to imaginary space, that it may be understood by reason, as well as perceived by sense'. A body is actually defined as 'that, which having no dependence upon our thought, is coincident or co-extended with some part of space'.¹⁰ Yet even this definition maintains his stress upon the ontological distinction between the real extension or magnitude of objects and the 'imaginary space' which depends 'upon our cogitation'.

Hobbes' definition of *place* as the space 'coincident with the magnitude of any body' corresponds roughly to Descartes' conception of 'internal place'. The Aristotelian relational definition (corresponding to Cartesian 'external place') is

rejected on the excellent ground that, since every part of a body occupies a *part* of its place, its place cannot be constituted by what surrounds it. Relational answers to the question 'where?', such as 'in the country' or 'in the chamber', are simply more or less determinate indications comparable to pointings, and are not direct accounts or definitions of the place of an object. Place is always conceived of as immovable, and motion is defined in terms of change of place.¹¹ On the nature of time, however, Hobbes seems to have been ready to stay fairly close to Aristotle: 'Time is the phantasm of before and after in motion.' The notion of *before* and *after* is that 'of *succession* in the motion of a body, in as much as it is first *here* and then *there*'. The reason for this definition is that the time must be measured by some motion, for we can represent and quantify time only by supposing uniform motion on a line, such as that of a shadow on a sundial. Yet although time is an abstraction from motion, and motion is the measure of time, the definition of time as the measure of motion is rejected.¹²

Hobbes' position combines apparent incompatibles. He insisted, like both Aristotle and Descartes, that space and time, conceived of as quasi-containers within which things exist and move, are mind-dependent 'phantasms' or abstractions from what does exist in its own right. On the other hand he found these philosophers' conveniently consonant doctrines about place and the vacuum quite unsatisfactory, and employed the notion of space not only in his definition of motion, but in his definition of body itself. He held, it seems, that space must enter into any account of reality, but that it is not a part of reality itself simply because it is nothing. Yet in having to allow it quantity we have to treat it as if it were something, a sort of shadow of body.

Gassendi took a different line. Faced with the same question, whether space is something or nothing, he plumped for 'something', and something distinct from body for all the usual reasons. Like Hobbes he rejected the relational account of place, supposing that even God could move the entire universe as a whole in space, as allegedly He could have created it a thousand years earlier than He did. The vacuum, both around and within the world, is not only possible but actual. Space 'cannot act or suffer anything', having the merely negative attribute of allowing things to occupy or pass through it. To the theological objection that, if space is a positive thing, then something exists uncreated by God, Gassendi responded unrepentantly that the 'imaginary space' of orthodox doctrine must in any case be so regarded. For 'imaginary' cannot mean that it 'depends on the imagination, like the chimera', but simply that we 'have an image of its dimensions by analogy with the dimensions that appear to our senses'. Besides, he added, it is orthodox to hold that even such physical and substantial things as essences are eternal and uncreated.

Time is allowed similar reality. There is an 'incorporeal duration independent of bodies', which extends beyond any beginning or possible end of the universe. Although time is measured by motion, it does not 'depend on motion or its parts, whether numbered or not'. There may be indefinitely many motions, but only one time. Like the comparable distinction with respect to space an ontological

distinction between 'real' and 'imaginary' time (a distinction which roughly corresponds to Descartes' distinction between 'time' and 'duration') is rejected. Gassendi's discussion was no doubt a model for Locke's in its elaborate comparison of space with time. As space is immobile, time flows at a uniform rate. As space (unlike body) is invisible, time (unlike physical processes) cannot be stopped. The universe, and every individual thing, fills a particular portion of space and of time. God exists everywhere and always. The dimensions of space correspond to the length, breadth and width of bodies, the dimensions of time to their motions. Time is no more dependent on motion or posterior to it, than extension is dependent on, or posterior to, yardsticks. But perhaps Gassendi's most striking doctrine about space and time is that they are 'incorporeal natures' which transcend the supposedly comprehensive dichotomy of substance and accident. Hence 'all being is substance or accident, or place, in which all substances and all accidents exist, or time, in which all substances and all accidents endure'.¹³

Henry More added to the realist argument on behalf of an immovable, invisible extension which is 'actual, necessary and independent on our imagination'. He employed a graphic example to refute Descartes' definition of movement: a suitably inclined coloured line within a rapidly spinning, solid, transparent cylinder will describe a visible cone, not in matter, for it is motionless in relation to the surrounding matter, but in space. Moreover, the space a body occupies cannot be, as Descartes had said, like a genus, since individuals and species cannot leave their genera behind. Nor is an empty space, as Hobbes had suggested, the possibility of a body, since, if it were, then the same space full could not be distinguished from the body which fills it, as actuality cannot be distinguished from actualized possibility. More's aim was to establish that space is an incorporeal and, indeed, spiritual entity, and to refute 'that prejudice, that whatever is extended, must be matter'. The properties of space, including its priority to all created things and to their motion, are such that it can only be 'the holy essence of God'. It has a property of all spirit in being 'penetrable' or capable of overlapping other substances material and spiritual. More had a curious theory that finite spirits, like all particular things, are extended, but that they differ from matter in their 'essential spissitude', a capacity to vary their extent by doubling back on themselves, as it were in a fourth dimension. Every spirit is where it acts, and God exists and acts everywhere. Like Malebranche and Berkeley after him, More offered his theory in interpretation of the text, 'In Him we live, move, and have our being'.¹⁴

Although we know that Locke had paid some attention to Gassendi's views on the question before 1671, the two drafts of the *Essay* written in that year (apart from passing mention of Descartes' dogmatic denial of a vacuum)¹⁵ avoid direct discussion of the ontological status of space and time. Instead, *Draft A* subordinated discussion of extension and duration to the argument that our idea of infinity is constructed and negative. In *Draft B* the whole discussion is more systematically located in a general discussion of relations, while a long account

of measurement, although leading into the topic of infinity and, more particularly, eternity, is nevertheless potentially free-standing. There is also a discussion of our idea of place.¹⁶ A number of themes to become prominent in the *Essay* were thus sketched out long before it achieved its final form, and it is perhaps appropriate to consider how they were developed in the published work before examining Locke's views on the largely separable ontological question. Not that these themes had no possible connection with ontology, but the only implications to be spelt out in the earlier writings were epistemological.

The chapter on the simple modes of space opens with a constructivist explanation of ideas of size and shape which is rather like that of Hobbes. The purpose, in effect, was to rebut an ancient argument for innate geometrical concepts and knowledge which was based on our ability to reason about figures with dimensions other than those which have fallen within our experience.¹⁷ On Locke's account, the ideas of such figures have all been constructed out of elements of which we have had experience. Since the elements employed are all of the same type, the construction is called a 'simple' rather than 'mixed' mode. It has already been remarked above that Locke can here be accused of having combined dissonant models in explanation of what it is to be an idea of a simple mode. He seems to have recognized (and yet it may seem difficult to reconcile) the conflicting claims to ultimate or elemental simplicity of, on the one hand, determinables such as 'distance', 'capacity' or 'extension', and on the other hand of determinate elements or units such as the 'sensible points' proposed at *Essay* II.xv.9. Yet Locke probably saw the two models as complementary. We have experience of lines and figures having determinate lengths, sizes and shapes, and can construct other determinate ideas falling under the same determinable by division and combination of these given elements. The analogy with numbers is overt and reflects the extent to which these first edition chapters retained the function of the discussion in *Draft A* of leading into the crucial topic of infinity. Ideas of determinates of the determinable *number* can be multiplied indefinitely by the addition of further elemental *units*. Only by such aggregation can the determinable be made determinate, so that both models for 'simple modes' apply together.

Another natural point of objection to this opening passage concerns measurement. Locke asserted that 'Men, for the use, and by the custom of measuring, settle in their Minds the *Ideas* of certain stated lengths, such as are an Inch, Foot, Yard.' Such ideas can serve as the elements from which we can construct an idea of any distance whatsoever.¹⁸ Both Leibniz and Berkeley protested that there cannot be a distinct idea corresponding to each unit of length, although the form of their objections was very different. For Berkeley there cannot be a single sensory idea corresponding to each objective 'stated length' because sense impressions must be supposed to vary, not only with the distance between two external perceived points, but with the distances of the points from the eye, with the size of the perceiver and so forth. An inch nearby may correspond visually to a foot at a greater distance. The objection depends for its

full force on the presupposition to which Locke seems to have been theoretically committed, that the idea must appropriately ‘resemble’ or at least ‘represent’ the corresponding objective length, as he understood those terms. Yet how can what is supposed quantitatively invariable ‘conform’ to a quantitative variety?¹⁹

We can hardly defend Locke here by means of an unLockean conception of the representative relationship, by stressing that we see three-dimensionally, so that a nearby inch does not after all characteristically look like a distant foot. If such considerations *are* allowed to enter the argument, however, the issue becomes in part transformed into the interesting question as to how far determinate length and area can be the objects of *perceptual* knowledge. First, it is clear that we can often simply see or feel quantitative relations between things, so that it makes sense to talk of strictly sensory illusions as to the relative length of two lines. The famous Müller-Lyer illusion, in which equal lines appear to be different lengths, is just such an illusion. The problem is whether we can take as literally our talk of seeing, or telling by eye, just how far it is from A to B. For it seems that precise quantitative estimates, however dependable in any particular case, should not be included within the scope of perceptual knowledge. The workman who cuts 40cm lengths ‘by eye’ does not see anything that the rest of us fail to see. His capacity is perhaps rather a recognitional than a perceptual one. Being 40cm long is not a sensible quality.²⁰

For Berkeley, however, what is problematic is not the measurement of the ‘internal’ length or idea, but the measurement of the realist’s supposedly sense-independent, external length. The latter, since it is supposed to resemble each of an indefinite variety of particular ideas as it is seen from different distances, cannot be ascribed a determinate quantity. Each idea, he held, has its size precisely determined by the number of ‘minima’ that compose it. The postulated independent object, however, has to be regarded as infinitely divisible and so in itself indeterminate in size.²¹ Berkeley’s conclusion was that the notion of an independent object is absurd, but his approach, ingenious as it was, sets the truth about measurement on its head. The notion of determinate distances and shapes (or even of precise spatial relations) on an immediately perceived visual or tactual field will not survive examination. When two lines of measurably the same length *look* different lengths, there are not two internal lines in the mind of the perceiver which *are* measurably different lengths.²² Look up at the centre of a rectangular ceiling, and no determinate answer can be given to the phenomenological question as to the precise shape of the area on the ‘visual field’ taken up by the image of the ceiling, or to the question whether opposite edges of that area are parallel.²³ It is not belief in the determinate object, but Berkeley’s belief in the measurable, determinate ‘idea’ which is the philosophical illusion.

Leibniz, on the other hand, evinced a rather better understanding of the relationship between real length and our ideas of it when he objected to Locke that ‘no-one can say or grasp in his mind what an inch or a foot is’. His point was that ‘the significance of these terms can be retained only by means of real standards of measure which are assumed to be unchanging, through which they

can always be re-established'.²⁴ As we shall see, it is probable that Locke would have responded to this profound criticism by arguing that the assumption of unchanging physical standards itself could have arisen only because certain objects *seem* not to change their length over time: in the end (or, rather, at first) we must judge according to appearance, i.e. by our 'ideas' of length and equality.

The view that judgements of extent are relative to determinate ideas of extent was not in any way peculiar to Locke, and can be found, for example, in the sceptical arguments of Arnauld, Malebranche and Bayle. In all three it is linked with arguments relating to infinite divisibility, and in Bayle it is also linked to a supposed analogy with secondary qualities.²⁵ To Berkeley the view would have seemed a universally orthodox doctrine fraught with well-known paradoxes. In Locke's account, however, scepticism gets a foothold only in connection with the parallel view of the measurement of time. No doubt Leibniz's response to him had been sharpened in advance by reflection on the French philosophers. Nevertheless the response was appropriate. A tribe which did not employ physical methods of measurement might make such vague or, if not vague, relative perceptual judgements of distance and length as 'that is near', 'these are equal', 'that is further off than this' but could have no terms corresponding to 'yard' or 'metre'. Admittedly we can estimate a distance in metres without measuring it physically, and that estimate may even be composite in so far as we may mentally mark off the distance metre by metre. But all such estimates depend for their meaning on the possibility of real physical measurement. On the other hand, that possibility obviously depends on the prior possibility of some immediate quantitative judgements or other. The point is that the latter will not be numerical, or involve such units as *metres*.

Similar issues arise rather more clearly in Locke's chapter on duration. Like the otherwise different theories of Descartes and Gassendi, Locke's account distinguishes the idea of duration in general, which we acquire from experienced succession, from the idea of time or measured duration. He believed that time presents a special sceptical problem because we 'cannot keep by us any standing unvarying measure of Duration' as we can of extension. We cannot bring together 'two Portions of Succession' in order to determine their equality. We must therefore rely on the supposition that certain recurrent phenomena, whether natural motions or other sorts of change ('the freezing of Water, or the blowing of a Plant'), divide up duration into equal parts. The succession of ideas itself constitutes the primitive clock through which we must perceive and judge any other succession, and which gives us our sense of time. Of the 'seeming equality' of periods defined by such objective recurrences as the rotation of the earth we initially 'have no other measure, but such as the train of our own Ideas have lodged in our Memories'. Thus there is a natural unit of measurement, the period from one idea to the next, which Locke took to 'vary not very much in a waking Man', and which he proposed to call an 'instant'. Measurement by 'instants' is provisional, however, for, like any experienced succession, the

succession of ideas is liable to be less than perfectly regular, and is in any case interrupted. Whatever regularity we employ as a unit, we presume that its cause 'which is unknown to us, shall always operate equally'. In other words, Locke recognized that behind our determinate judgements of time lies the presumption of an unchanging, enduring nature and natural law. The closer our clocks come to the most fundamental natural regularities, the better they are.²⁶

Berkeley, as we should expect, denied objectively determinable duration and pronounced the succession of ideas to be the only determinate measurement of time. Each spirit therefore has its own internal time.²⁷ Leibniz just as predictably rejected the notion of the primitive internal clock, although on the apparently contingent ground that our 'perceptions never provide a sufficiently constant and regular train to correspond to the passage of time'.²⁸ Yet the very conception of a succession of 'ideas', regarded as discrete and countable without regard to what they are ideas of, is unacceptable, so that the primitive sense of time, like the sense of distance, is necessarily indeterminate. Objective clocks are necessary for the determination of time, even if we must rely on a pre-horological sense of time for the initial selection of our clocks.

Leibniz's commentary on Locke's discussion supplied, we can assume, an important part of the material transformed by Kant in his deduction of an objective, permanent natural world from the necessary conditions of time-determination, the famous 'Refutation of Idealism'.²⁹ Yet the argument can be extended, since the analogy between time and space is closer than Locke suggested. Measurement of time is not significantly the more vulnerable to scepticism since, as Leibniz pointed out, we cannot assume without question that any particular yardstick has remained the same length over time. In the measurement of time we can determine our unit either by such relatively dependable particulars as the sun, or by a universal relationship or law such as the one Locke mentioned, the time of the swing of a pendulum having a certain length. Similarly, in the measurement of space we can employ either such relatively dependable particulars as the standard metre in Paris or else, more directly, a law. We could even, as Leibniz proposed, employ that very same law (assuming that we can otherwise determine the time the pendulum takes to swing, rather than that we can otherwise determine its length). Whichever method we choose with respect to each modality, we presuppose the uniformity of nature. Without that presupposition the claim that an object has remained the same length, or that two successive periods have been equal, would lack truth-conditions. Consequently, as Kant argued, if whatever exists in space and time necessarily does so determinately, then it is necessary that whatever exists should be subject to law. Since the former does seem necessary, that is an implication of great importance. Yet, if it did not occur to Locke to draw it even with respect to time, that might have been for more than one reason. First, the principle of universal causality was for him an intuitive truth presupposed in the discussion of time, not looking for support from it. It was only with scepticism like Hume's that a need for Kant's argument arose. Second, the link between determinate

duration and law was for Locke less than absolute. The distinction between duration and time allows, he thought, for a determinate duration which is conceivable but only hypothetically measurable: the duration of God before the creation of nature and time and after their joint annihilation. Yet such a distinction is untenable precisely because nature supplies the truth-conditions of any determinate temporal judgement, not just our grounds for making such a judgement.

This is perhaps an appropriate point at which to return to the topic of the ontological status of space and time. Some early thoughts on the question are recorded in Locke's journals for 1676, and seem to have been firmly Hobbesian. Later entries over nearly two years elaborated the same themes. Space is nothing real, but the 'bare possibility of body to exist'; or is *nuda relatio*, having no actual existence apart from the dimensions of bodies in relation. Its merely abstract character explains why it is immovable and indivisible even in thought, although it is commensurable with what is divisible. Those who regard space as a positive thing have been misled by the way we talk. There was one minor hypothetical concession to the Cambridge Platonists. If it is after all impossible to conceive of pure nothing, then empty space must 'belong to the being of the deity'. Two epistemological features of Hobbes' doctrine were taken to be unquestionable: the idea of space is abstracted from our experience of bodies, and the idea of a vacuum makes sense.³⁰

Despite the prominence of these last two principles in the chapter on the modes of space in the *Essay*, the general tone of the discussion had by then become considerably less influenced by Hobbes, and indeed ambivalent or agnostic as between Locke's earlier Hobbesian views and the 'Cambridge' theory espoused by Newton:

whether any one will take Space to be only a relation resulting from the Existence of other Beings at a distance; or whether they will think the Words...of the inspired Philosopher St Paul, *In Him we live, move, and have our Being*, are to be understood in a literal sence, I leave every one to consider.³¹

The issue seems to have been unimportant to Locke compared with the anti-Cartesian thesis, advanced by More as well as Hobbes, of the intelligibility of a physical vacuum, whether within the material world or beyond its bounds. Moreover Locke's argument for this conclusion was no longer cast, as it had been in the journals, in Hobbesian terms. The possibility that someone 'at the extremity of corporeal Beings' should stretch out a hand he now, it seems, explained by the prior existence of empty space, rather than *vice versa*; and there was no longer talk of 'imaginary space'. The answer to the question whether space is something or nothing was now, at least implicitly, 'something'. An agnostic version of Gassendi's argument was employed in response to the question whether that something is substance or accident: the time to answer that

conundrum is when we understand what are substances and what accidents, and whether everything must be one or the other. Like More, Locke proposed to distinguish verbally between the expansion of space and the extension of body: between ‘expansion’ (for More, ‘extension’) and ‘extension’ (for More, ‘amplitude’).³² The attributes of space which distinguish it from body, in effect taken over from More, seem to have been understood realistically, rather than as a mere consequence of abstraction: ‘its Parts are inseparable, immovable, and without resistance to the Motion of Body’.³³ In II.xv, moreover, the chapter in which ‘expansion’ and ‘duration’ are compared, Locke seems to have been unreservedly realist. He explicitly rejected the term ‘imaginary space’, for example, attributing its use to an unnecessary reluctance to allow infinite extension to God as well as infinite duration.³⁴

There is, however, one striking concession which the argument of II.xiii appears to make to the anti-realist tradition, for it seems to endorse the purely relational account of *place* given in *Draft B*, but which Hobbes himself rejected. Consequently ‘we can have no Idea of the Place of the Universe, though we can of all the parts of it’. With apparent irony, Locke remarked that someone who formed a clear and distinct idea of the place of the universe could ‘tell us, whether it moves or stands still in the undistinguishable *Inane* of infinite Space’. However, ‘the word Place, has sometimes a more confused Sense, and stands for that Space, which any Body takes up; and so the Universe is in a Place’.³⁵ This last remark has been interpreted (wrongly, as we shall see) as a dismissive allusion to the view of Hobbes and More which Newton re-expressed in his general definition of place as ‘a part of space which a body takes up’.³⁶

Two questions can be asked about all this which cast some light on each other. First, how, if at all, can we reconcile the realist and anti-realist elements which seem to exist together in the account of space in the *Essay*? Second, what effect, if any, had the reading of Newton’s *Principia* on Locke’s doctrine? Although we cannot expect to answer the second question directly, we can appeal to some striking differences between the extant draft of 1685 (*Draft C*, written a year before the publication of *Principia* and still unpublished) and the first edition itself. Admittedly the chief agnostic passages of II.xiii were already present in *Draft C*, including the Gassendist response to the question whether space is something or nothing. There was also the more positive suggestion, corresponding to that of *Essay* II.xv.2, that God is coextensive with infinite space. Yet the anti-realist elements in the argument of *Draft C* were far stronger than in the *Essay*. The traditional term ‘imaginary space’ was freely used, whereas in the *Essay* it was explicitly criticized. Moreover the only idea of place to be identified in *Draft C* was the relational one, explained in terms which impugn the reality of infinite space: ‘place and time are real relations resulting from the existence of real different Beings’. They are identifiable ‘by real and distinguishable marks and fixed boundarys and so seeme to be determinate portions of those infinite Extensions of space and duration we have in our imaginations’. Someone who ‘can finde out and frame in his minde clearly and

distinctly where the place of the universe is' will know 'that place is not a relation of real distinct beings without its bounds but something else but till then I thinke must be content with me to take it for a Relation'.³⁷

All this was considerably toned down in the *Essay*, a point which raises the question of what exactly was meant there by the added reference to a second, 'confused' sense of the word 'place'.³⁸ Because the explanation of this 'sense' is comparable to Locke's explanation in other contexts of the idea of the *extension* of a body,³⁹ it might seem that he was complaining of the merely verbal confusion of calling a body's extension its 'place', as in Descartes' expression 'internal place'. In that case he would have been stressing that the only place which a body can occupy temporarily and leave behind is its relational or 'external' place. But another and, as I think, much preferable interpretation grants the addition more immediate relevance to its context. This reading takes 'confused' to be contrasted with the immediately preceding, ironical supposition of a 'clear and distinct' idea of the place of the universe in absolute space, so that the 'confused' idea in question is that of the absolute place of a body, the portion of absolute space which it at any instant occupies. In that case Locke must be seen, not as having dismissed the Newtonian view, but as having made a concession to it. For the context would suggest that such an idea is 'confused' in much the same way as the indispensable notion of substance is confused. Absolute place, like substance, must in general be supposed to exist, but it is not something which we are capable of determining or distinguishing or re-identifying from one moment to the next: or of which, therefore, we have a 'clear and distinct' idea. On this interpretation, the reason for the impossibility of determining the absolute place of a body, and so for the 'confusion' of the idea, is not that the general notion of real, absolute space is just a muddle as the anti-realists believed (and as Locke himself had previously believed), but simply that 'those uniform infinite Oceans of Duration and Space' are of their nature featureless.⁴⁰ On the everyday, clear and distinct notion of place, on the other hand, which is acquired and employed on the basis of experience, a body's place is simply its relation to other bodies, just as the date of an event consists in its relation to another event. So although Locke did not bring both notions under Newton's general definition of place, the Newtonian distinction between absolute, real but epistemically inaccessible place and relative, sensible place seems here to be explicit in the *Essay*.

It is, then, difficult not to read *Draft C* as transitional between the Hobbesian view of the journals and the firmly 'Cambridge' doctrine (if overlaid by Locke's principled agnosticism) to be found in the *Essay*. It also seems permissible to speculate that Locke was nudged along the realist road by his reading of Newton, although he could have been motivated by considerations of consistency alone. Perhaps Newton's own most striking contribution to the debate, however, was his claim that the question as to whether a body moves in absolute space can at least sometimes be determined from the 'causes and effects' of absolute motion. The absolute circular spin of two globes connected by a cord might be

deduced from the tension of the cord, 'even in an immense vacuum, where there was nothing external or sensible with which the globes could be compared'.⁴¹ The point at best falls short of a recipe for determining absolute position in 'those boundless Oceans of Eternity and Immensity', but if Locke did have Newton firmly in mind it is a little surprising that he made no allusion to the famous thought-experiments.

I have little to say on the puzzling philosophical problem of the ontological status of space and time. The theory of relativity may have changed the shape of the issues, but seems not to have resolved them—least of all in the direction which might be suggested by the word 'relativity'. Realist and anti-realist still find grounds for dispute, and some of the more impressive grounds for realism are new. Perhaps the link between an anti-realist view of space and a more general idealism, a connection which began to emerge in Locke's time in the arguments of Leibniz, Bayle and Berkeley, points to the necessity that a consistent physical realist should also acquiesce in the reality of space-time. Or perhaps the issue of the 'reality' of space is not one question but several, including the question whether, as one explanatory model requires, we should ascribe natural properties to space as well as to things in space in the interpretation of relativity. Perhaps one positive contribution of the theory of relativity to the resolution of the seventeenth-century issues is that it now seems possible to make sense of the proposition that space is finite.

Thought about particulars

The above analysis of Locke's conceptions of space, time and place may help us to grasp their role for him with respect to particulars and our ideas of particulars. For the *Essay* followed one piece of Gassendist and 'Cambridge' doctrine unswervingly: every finite thing exists in space and time, as God exists throughout them. If there are finite immaterial substances, the principle emphatically applies to them too, whatever difficulties we may find in conceiving of their place-occupancy. A spirit must be where it acts. The identity of a spirit over time, as well as that of body, is determined by its continuous relationship to 'its determinate time and place of beginning to exist', i.e. by its spatiotemporal continuity.⁴²

This doctrine supplied Locke with his explanation of the relationship between particulars and universals. The latter, on a traditional view, exist outside space and time, a metaphor which Locke reinterpreted in effect as the thesis that, while to think of a particular existent is to think of it in its spatio-temporal location, to think of a universal through the particular is simply to abstract from that location. Discussion of abstraction will be postponed to the chapters following. In the present one I will consider what is involved in Lockean ideas of particulars.

It could be said that Locke gave no serious theoretical consideration to the question of the particular reference of thought. He seems simply to have assumed that the work can be done by the qualitative conformity of the idea to the object, together with the inclusion in the idea of the 'circumstances of real Existence, as Time, Place, or any other concomitant Ideas'.⁴³ Obviously, special weight would need to be assigned to time and place, since the inclusion of concomitant qualitative ideas would serve only to make the idea in question *complex*, rather than *particular*. Yet the incapacity of ideas of time and place by themselves to particularize the idea of an object is perhaps indicated by Locke's own account of such ideas in an argument which certainly *is* theoretical. For he held, as we have seen, that our ideas of time and place are relational. Yet if we try to explain what it is for a thought to be about X in terms of a thought relation to Y, we leave ourselves with the task of explaining thought about Y. It seems that we

need to have thoughts about particulars *before* we can have ideas of time and place. Locke's explanations therefore seem circular.

From explanations of the reference of thought suggested in the present work, it might seem that Locke's whole approach to thought of particulars is misguided. As we have seen, what links a thought to a particular individual as its object is not just the intrinsic content of the thought but a certain sort of causal link between the individual and the thought: in the most straightforward cases a link supplied by the mechanisms of perception and memory. Locke would doubtless have agreed that in order to think of an object experienced in the past one has to remember it, while the ideas of particulars in his examples are taken to be formed in actual sense-perception. Yet he laid no theoretical stress on the origin of the idea in experience in his explanation of what makes it an idea of one particular object. There seems to be no conscious intimation in the *Essay* of the insight so much emphasized by modern writers, that one could not look into someone's mind and from that alone, together with knowledge of the things in the world, know which particulars he or she is thinking about. If my idea of Aristotle were compared with that of someone who just once saw Aristotle as a baby, only such very general attributes as membership of the human species would be likely to be common to both. Moreover my conception of Aristotle might be extremely vague or even erroneous, and yet still be a conception of Aristotle. In the extreme case of traditional reference, we are able to think about an individual while knowing virtually nothing about it (indeed, while having almost entirely false beliefs about it), just because we have taken over the referential use of a proper name from others. For perhaps the last two millennia or so, that has been true of thought about Jonah. 'Who was Jonah?' is, after all, (at least, if Jonah existed) a question about Jonah, even if no one can now answer it very well.

Not that the relation of reference is a purely causal one. As we have seen, some conformity between thought and object is requisite too. For it to be true that I see this crumpled shirt as a face, then, to put it roughly, the parts of the face must appear where the parts of the shirt are disposed. A merely causal connection in the nervous system between sensory input from the shirt and the vision of a face is not enough. Analogously, I cannot pick up Aristotle's name from others' usage and use it myself to refer to him, if I believe it to be the name of a girl living in Streatham. Even Jonah I know of as an ancient Hebrew about whom a certain story is told in the Bible. But what the role of causality in reference might suggest is that thought about particulars can be understood quite independently of thought about time and place, so that the circle which seems to be formed when we bring Locke's accounts of the two sorts of thought together can easily be avoided: reference has nothing in particular to do with knowledge of where and when a thing existed.

Such a solution would be both too easy and unjust to Locke, since it depends on the assumption that understanding reference is a matter of understanding the essence of reference present in each and every case of it. The modern 'causal theory of reference' is as one-sided as the causal theory of knowledge. As it was

argued in [chapter 15](#) above, in order to understand knowledge in general it is necessary to understand primary knowledge in particular, and how secondary knowledge would be impossible without primary knowledge. The same seems true of reference. Indeed a category of primary reference has in effect already been proposed above, when it was suggested in [chapter 21](#) that the spatio-causal relationship between object and sensation is itself presented in sensation. That is to say, the very way in which the object is presented renders perspicuous the extrinsic referential relationship. The perspicuity of the causal relation is not to be disentangled from that of the spatial relation: the object strikes the eye from a direction, its pressure is felt at a place. A directional illusion, as in a mirror or a mirage, is *ipso facto* an illusion as to how the object is acting on me. Something analogous seems true of time and primary memory, in which I am aware of the origin of my remembering in the past event remembered. The temporal relation, together with the causal relation, enters into the content of the remembering. The event is recalled as past experience, and I know how it is that I am in a position to think about it. Thus the role of time and place in Locke's account of ideas of particulars might not unreasonably be attributed to his having had some intimations of their role in the perspicuous primary reference of perception and memory.

If we took Locke's philosophical field of vision to be, to that extent, filled by the basic case of the reference of our perceptual states, then we might be inclined to speculate that he would have responded to the charge of circularity by proposing that the relative place and time of things can always in the end be tied down to the here and now of oneself as subject or perceiver. Yet there is nothing in his discussion of relative time and place which gives the least hint of an intention to bring us back to demonstrative reference, or to find a frame of reference in the subject himself or in the here and now of the thought in question.⁴⁴ For the seemingly circular principle that place and time are always given in relation to 'certain known points fixed in distinguishable sensible things'⁴⁵ is illustrated by the use of such 'landmarks' as the birth of Christ, far from the experience of the implicitly postulated subject. Locke meant simply that co-ordinates relate to convenient public reference points selected from all that is known. Thus the model or principle underlying both the discussion of place and time and the remarks about ideas of particulars (let us call it 'the principle of mapping') was roughly as follows. In order to get to determinate grips with the 'boundless invariable Oceans of Duration and Expansion' we must build up a picture or map of known objects in their spatio-temporal relations. To fit anything new into the map, we must measure its distance from something already there. But our very capacity to think of particular objects in their absence depends on our possession of such a map, i.e. on our capacity to think of things through their position on the map, within a system of relations.

Despite all the objections which can be brought against Lockean ideas of particulars, the principle of mapping is in the end extremely plausible. It is, of course, as we have seen, easy to find apparent counter-examples, examples of

our thinking of an object in its absence without being able to locate it in any way in space and by time. For example, we may think of something without remembering when or where we saw it, or whether we saw it or only heard about it, or even whether the object was reported or predicted. Yet such cases are not primary cases of thinking of something, just as our remembering an event while supposing that we are merely imagining it is not a primary case of remembering an event. Unless *in general* we apprehended the spatio-temporal relations which the objects of our thought bear to one another and to ourselves, there could be no objects of thought of which that is not true. One argument for such a conclusion lies, as in the case of perception and memory, in the impossibility of separating awareness of causal relations from all awareness of spatio-temporal relations. *Primary* or *perspicuous* reference involves some awareness of both, whether immediate and demonstrative or dependent on our possession of some sort of a mental map.

Talk of a mental map of reality must, of course, be understood metaphorically and dispositionally, but that does not mean behaviouristically, or that the dispositions are not actualized in consciousness. Not only the sophisticated capacity mentally to recall or construct a map of Oxford in the literal sense, but the simple capacity to visualize two houses in spatial relation, the more interesting capacity to visualize the table behind my back in relation to my body (when 'image' and bodily sensation combine as elements of a single field), and even the capacity, as I sit in my room, to 'feel' the direction of the Sheldonian Theatre (perhaps there is a sort of motor image of a step towards it, or an actual bodily intimation of direction) may each of them be a capacity of the general type without which it would be impossible to think of currently unperceived objects at all. Not that the impact of the 'map' on consciousness must necessarily be supposed to take the form of 'images', but the time for that argument is not quite yet. Consideration of the extent of the content of consciousness and the forms of conscious representation is better postponed until after an exposition and criticism of Locke's imagist theory of abstract thought.

Locke's theory of universal knowledge in context

Locke's is probably the earliest theory of abstraction still to receive critical discussion as a part of the normal philosophy curriculum in universities in the English-speaking world. Yet, as disputes about its interpretation and vulnerability to Berkeley's famous criticisms bear witness, it is easy to miss the meaning of the theory and its point unless it is located in the long-standing context of debate. In fact, the notion of abstraction was common property in seventeenth-century philosophy. It was associated in Aristotelian philosophy with the notion of a distinction of thought or reason, such as the logical distinctions between form and matter, between species and genus, and between substance and accident. More relevant to our present concerns, however, is its connection with the doctrine that the active intellect forms universal 'intelligible species' from particular images in the imagination, images deriving in turn from 'sensible species' intromitted in sense experience.

Today it is easy to think of a doctrine of abstract ideas as above all an element in empiricist theories of concept-acquisition, but the notion was just as closely and naturally associated with the Cartesian conception of the methodical analysis of the complex into the simple. As abstraction was explained in the Port Royal *Logic*, 'because of the limited scope of our minds, we are unable to understand perfectly even things which are only slightly complex unless we can consider these things part by part or with respect to their different aspects'. In its proper sense, according to the *Logic*, 'abstraction' is not the distinction of parts which can actually stand alone, but the distinction of inseparable aspects, as 'when we consider a mode without paying attention to its substance', or attend to just one dimension of three-dimensional space. Hence geometers do not postulate the existence of lines without breadth, they simply 'consider length without paying attention to breadth'. By abstraction we can arrive at the determinable essences of substances, although such essences only actually exist with determinate modifications. If I attend equally to all the accidental characteristics of an equilateral triangle drawn on paper, then 'I shall have an idea of that particular triangle alone'. But if I focus on its being 'a figure bounded by three equal lines', then 'my idea will represent all other equilateral triangles'. If I disregard the equality of the sides, then it will represent all triangles whatsoever. 'Thus I can

ascend by degrees to an idea of extension itself.' The discussion prepares for the conclusion that abstraction explains the difference between universal and particular ideas, between those ideas which represent many objects, and those which represent just one object. But it is also taken to explain an Aristotelian doctrine of some importance (as it will be found in Volume II) in the interpretation of Locke. That is the doctrine of predicables, which draws the distinctions between species, genus, difference, property and accident. The fact that we can draw these distinctions in thought does not show in any particular case in which we apply them that we have grasped what the scientists need to know, 'the true genera of things, the true species of each genus, the true differences of species, the true properties'.⁴⁶

The Port Royal discussion is typically Cartesian, reorganizing and elaborating a passage from Descartes' *Principles* in which, after claiming that 'all universals are simply modes of thought', he had gone on to suggest that the employment of the predicables is similarly mind-dependent, and to examine the variety of real and logical distinctions.⁴⁷ It is clear that all such passages should be interpreted in a way compatible with the doctrine that the simple end-products of abstraction are characteristically purely intellectual and, what is more, innate ideas. Arnauld himself made just that point explicitly:

the perception of a particular body, only to be had through the senses, can awaken in us the idea of a body in general: e.g. a square drawn on paper awakens in us the universal idea of a square. Yet that does not prevent... the universal idea of a square from being a pure intellection, even though it is accompanied by an image in the brain. For our soul does not stop at what is particular, whether in the image in the brain or in the image which is drawn on the paper, but applies itself solely to the abstract idea of a square in general, which can be traced neither in the brain nor on the paper.⁴⁸

Moreover, Descartes' relegation of universals to the mind did not prevent him from holding that our true universal beliefs owe their truth to something lying outside human minds, the eternal truths, essences or archetypes in God's mind. It is this theory, Scholastic except in its account of what divine archetypes there are, which allowed him to be at once both a kind of conceptualist and a kind of Platonist.

In the hands of seventeenth-century empiricists or imagists, on the other hand, the notion of abstraction served as a weapon not only against Aristotelian genera and species but also against Cartesian innate intellectual notions. Their line of argument went back to ancient empiricism and the universal 'stored notions' or 'preconceptions' of the Epicureans. Derived from repeated sense-experience, these self-evident principles of judgement and classification fix the meaning of the terms we apply to things. It seems from available accounts that the Stoics developed the model into a complex, self-consciously reductive theory of

universals, in particular of the Aristotelian kinds of substances. ‘Ideas’ or ‘forms’ are nothing but our own concepts, generic impressions formed in experience for the purpose of language. Everything that exists is particular, and the concept or conceived universal is not a thing but a quasi-thing such that predications of which it is the subject do not obey the logical rules which hold when an existent thing is subject: for example, *man* is neither Greek nor non-Greek, although a particular man must be one or the other; *man* can be in both Athens and Megara, but a particular man in Athens cannot be in Megara. At the same time a genus ‘is a collection of a plurality of inseparable concepts.... For [*animal*] embraces all the distinct animals’. Another reductive point was that the definitions embodied in these figments imprinted or inscribed on the soul are equivalent to hypotheticals: ‘Man is a rational mortal animal’ says only that, *if* something is a man, *then* it is a rational mortal animal.⁴⁹

One feature of this type of theory, mentioned in [chapter 1](#) above, is its (not particularly unreasonable) tendency to conflate acquiring primitive sensory concepts with acquiring evident criterial or definitional truths, and it is consonant with that tendency that Gassendi should have held that all evidence and certainty derives from the evidence of sense. As it was remarked in the same chapter, Locke seems at first, in *Essays on the Law of Nature*, to have adopted just that ambivalent conception of *praecognita* acquired in sense-experience as the starting-points of reasoning and science. In [chapter 2](#), moreover, it was noted that even as late as *Draft A* he proposed that our ideas of substances are at least quasi-propositional. Yet in the course of writing *Draft A* he effectively arrived at the *Essay*’s doctrine that, while ideas incapable of truth or falsity in a strict sense are derived from experience, definitions, like other *a priori* propositions, express a relation between ideas (in their case, identity). To that extent his later theory freed *a priori* knowledge from experience, whereas Gassendi’s seems to have been more faithful to the ancient models.

Those models might be suspected of suffering from another sort of ambivalence. On the one hand the universal conception is inscribed on the soul by ‘many memories of a similar kind’, and the genus contains a plurality of subordinate kinds. On the other hand, ‘the generic man is neither Greek (since then all specific men would be Greek) nor barbarian (for the same reason)’.⁵⁰ In other words, there is a question whether we are to think of the nature and formation of the inscribed concept through its extension or through its intension. Gassendi similarly held that there are two methods of concept-formation, and two sorts of universal idea. By the first method, ‘the mind as it were selects similar ideas and compresses them into one heap, which, so containing all of them, is made an idea of all together (*universarum*), and is therefore called universal, and general and common’. The second method is described as follows:

Although those particular ideas are in some respect alike or in agreement with one another (*mutuo convenient*), yet they also have many points of distinction (*discrimina*) by which they differ among themselves. The mind

therefore, viewing separately and so as it were drawing out of (*abstrahendo*) all of them what they all agree in, the points of distinction by which they differ having been removed or rather (*seu.*) not considered, keeps that abstractly considered object, containing (*habens*) nothing which is not common, as the common, universal, general idea, and it is just this which is called the genus.⁵¹

Gassendi seems in effect to be offering a psychological explanation of the possibility of thinking universally in either of two different ways, in terms of classes or in terms of attributes. Each sort of idea allows for the construction of a hierarchy of species and genera, and both are liable to related shortcomings. A 'universal' idea of *man* of the former sort which leaves out Americans is imperfect, as is a 'common' idea of *man* of the second sort which includes features not common to all men. Gassendi here remarks that it is impossible to imagine a man in general (*in commune*) so abstractly (*pure*) that he has no specific features, but we should keep in mind that man in general, as we want to consider him, ought to be free from all such points of distinction among men.⁵² It would seem that all universal ideas are necessarily imperfect in one way or other, but that we can make some allowance for their imperfection in thinking universally.

Gassendi brought his theory against Descartes' acceptance of eternal essences, although here he appealed only to the notion of a universal concept constructed by abstraction from actual men and serving as a rule for deciding what is to count as a man. 'Man is an animal' is an eternal truth, or true when no man exists, only in the sense that, whenever a man exists, he will be an animal. Similarly the universal triangle is not something real, with real properties: it possesses properties 'only in so far as the intellect, after inspecting...material triangles, has attributed such properties to it, only to give them back to the material triangles again in the course of demonstration'. How could there in any case be a real triangle without depth and composed of lines without breadth?⁵³ Despite the difference in its view of definitions, the *Essay* contains many passages which seem to echo Gassendi's arguments. Yet there is another philosopher in particular, with a rather more detailed explanatory model for universal thought, who seems likely to have had an important, perhaps decisive, influence on the form of Locke's own theory: namely Hobbes.

Hobbes' theory differed from the ancient and Gassendist accounts in that he denied that any special kind of idea is constructed by the mind which could be called 'universal'. On the contrary, 'there is nothing universal but names'. He did, of course, believe that in all thought-sensory ideas, images or 'phantasms' must be before the mind, in consciousness. These images occur in a more or less orderly train which comprises what he called the 'discourse of the mind'. The general principle of order or association among images is 'their first coherence or consequence at the time when they are produced by sense'. The most important type of such order, prior to language at least, arises from the

experience of cause and effect: 'as for example, because a man hath often seen offences followed by punishment, when he seeth an offence in present, he thinketh punishment to be consequent thereto' in the future. Things so linked in experience naturally 'signify' each other, as clouds are a natural sign of rain and rain of clouds. Expectations thus based on experience are manifestations of what Hobbes called prudence, a virtue to be found in animals as well as man.

What Hobbes called *science*, however, is impossible without general names joined in propositions. 'Names', i.e. words or expressions which can stand as subjects or predicates, are themselves experienced as sensations or images which are arbitrarily rather than naturally associated with other sensations and images, as the sound 'red' is conventionally associated with perceived and imagined red things. In virtue of this arbitrary association they function both as 'marks' and 'signs'. As marks they serve in internal discourse to record what we experience, in virtue of their capacity to revive or stimulate present images like past sensations or images. As signs they serve to make our thoughts public by stimulating in others images like our own. General or universal names are names which are thus associated in thought with the images of many things, rather than one thing, on the basis of the resemblance or equality between the things. General reasoning consists of a sequence of particular images accompanied by appropriate universal names combined in propositions. Truth and falsity are attributes of propositions, and a proposition is true when the predicate comprehends the subject, i.e. when whatever is named by the subject is named by the predicate, as in the proposition, 'Man is a living creature.' Such a truth as this one is 'eternal' just because it is equivalent to a hypothetical relating to language, 'if *man*, then *animal*'. Truth is 'evident' to the thinker when a certain relation exists, and is recognized, between the primary images and the concomitant names. The same kind of relation constitutes the 'evidence' of validity when propositions are combined in reasoning or syllogism.

One of the examples which Hobbes used in order to explain this relation is the materialist syllogism, 'Man is a living creature, a living creature is a body, therefore, man is a body'. Corresponding to the minor premise, 'Man is a living creature', we form, first, an image of a thing engaged in discoursing or reasoning, remembering that such an object is called 'man'; and, second, an image of the same thing moving itself, remembering that what is imagined is called 'living creature'. Corresponding to the major premise, 'A living creature is a body', we form an image of the same thing as was imagined moving, filling space, remembering that that is called 'body'. We remember that the same thing was imagined throughout, and so recognize that what is named by the subject of the conclusion 'Man is a body' is also named by the predicate, i.e. that the conclusion is true. Both verbal and non-verbal images are necessary for syllogistic thought because, Hobbes claimed, 'it is necessary to think not only of the thing, but also by turns to remember the divers names, which for divers considerations thereof are applied to the same'.⁵⁴

It is not entirely easy to get a clear view of the role of names in this account. They seem to be given a psychological role in the orderly stimulation of images, or in our being brought to focus on those aspects of the currently imagined object in virtue of which the name is associated with it. But *if so facto* names link the analysis of the features or aspects of a particular object of sense or imagination to other actual or possible objects. As Hobbes put another example, without language someone might recognize the equality of the angles of a particular triangle to two particular right angles, but the general rule will occur only to someone

that hath the use of words, when he observes, that such equality was consequent, not to the length of the sides, nor to any other particular thing in his triangle; but onely to this, that the sides were straight, and the angles three; and that that was all, for which he named it a Triangle.⁵⁵

The word 'triangle', that is to say, has the function of picking out, and enabling us to return to, that aspect of the particular triangle, its precise point of resemblance to other triangles, which is the ground of our use of the name. It therefore enables us to extend our conclusion to all triangles universally. This claim, like the account of syllogistic reasoning, is presented as a specific application of a general psychological theory about the 'consequence or train of imaginations'.

This kind of explanation of universal thought, as it will be seen, is deeply unsatisfactory, but at least Hobbes was facing up to the problem as to how consideration of an imaged object could be supposed to contribute to the apprehension of universal truth. He saw the pointlessness of postulating *queer* overlapping or indeterminate images of one kind or another, images necessarily inadequate for the task in hand, and he also had a clear sense of the oddity of supposing that one might syllogize without language. Yet the role he accorded to general names is psychological rather than logically essential. The arbitrary or conventional association of verbal with non-verbal images is supposed to free man from the tyranny of present experience and the relatively disorderly, easily disrupted natural association of non-verbal images. Hobbes' aim was to explain how language makes science possible, and to do so by means of a psychology according to which the same principle of association which explains animal prudence will, given the introduction of language, explain *a priori* rationality. Consequently there is no need to have recourse to a natural faculty of reason or pure intellect: 'for the understanding of the extent of a universal name, we need no other faculty but that of our imagination, by which we remember that such names bring sometimes one thing, sometimes another, into our mind'.⁵⁶ In effect, in inventing language man invented his own defining property.

Locke was neither an associationist nor, in the strict sense, a nominalist, but he was an imagist and shared Hobbes' need to explain universal thought in terms of the imagination. Since even that has been doubted, it is worth expounding his

arguments with some care. His first statement of his doctrine takes the simple idea of whiteness as its example. We make 'the particular Ideas received from particular Objects' general 'by considering them as they are in the Mind such Appearances, separate from all other Existences, and the circumstances of real Existence, as Time, Place, or any other concomitant Ideas'. The resultant 'precise, naked Appearances' become 'general Representatives of all of the same kind', membership of the kind belonging to 'whatever exists conformable to such abstract Ideas'. Thus the abstract idea has two linked functions, as 'representative' of a class, and as a 'standard' or 'pattern' to 'rank real Existences into sorts', as Locke put it. A 'real existence' is said to 'conform' to an antecedent abstract idea if it produces in sensation ideas or images precisely resembling the abstract idea.⁵⁷

Because of this second function as standard or pattern, some modern commentators have assumed that Locke was concerned with the question of how it is possible to recognize a recurrent feature or point of resemblance so as to apply a general term consistently. It is supposed that he tried to solve this problem by postulating a sort of permanent set of mental paradigms, like a colour-chart, which is compared with reality so that the right predicate can be selected. Yet if such a chart were necessary at all, then we should need another such chart to reassure us that the first one had not changed, or that the same chart had come to mind this time as last. In fact Locke simply and reasonably assumed the capacity to recognize features or points of resemblance. He was not concerned with any such 'problem' of classification, but with the possibility of universal knowledge, which he explained as follows.

Suppose that I perceive a relation between two particular ideas or images, A and B, functioning in my thought as abstract ideas. Since nothing will be counted as a member of either of the classes of objects *represented* by A and B unless it is capable of producing an idea in its abstracted aspect *precisely resembling* or conforming to A or B respectively, then whatever relation holds between A and B will hold between any member of the class represented by A and any member of the class represented by B. For example, it is Locke's view that, if we perceive that the angles of a particular representative triangle are equal to the angle on a particular representative straight line, then we *ipso facto* perceive with respect to the whole class of triangles (i.e. the class of whatever figures precisely resemble the representative in the abstracted respect) that every one of them has angles equal to the angle on any straight line. Thus the role of the abstract idea as 'standard' has simply to do with the logical determination of the class which it 'represents', and nothing whatsoever to do with a supposed problem as to how the mind is capable of recognizing recurrent features or points of precise resemblance. The abstract idea just is the phenomenal particular currently before the mind and representing all particulars which precisely resemble it in the respect upon which the mind is focused in abstraction. Or we could say that the abstract idea is the relevant *aspect* of the representative particular, so that it is unsurprising that Locke talked of finding abstract ideas 'in' perceived things:

'For what is known of such general *Ideas*, will be true of every particular thing in whom that Essence, i.e. that abstract *Idea* is to be found.'⁵⁸ The point of the theory is to explain how, although (as everyone agreed) the senses give nothing but particulars,⁵⁹ and although (as the imagists held) thought is bound to particular sensations and sensory images, yet we can 'perceive' relationships between universal classes.

It has been pointed out above (in [chapter 1](#)) that Locke's word 'idea' enjoys an innocuous and self-conscious 'type-token' ambiguity which makes possible 'occurrent' and 'dispositional' senses of the expression 'to have an idea'. The abstract ideas 'laid up' or 'set up' in the mind as 'standards' or 'patterns' and used as 'the Representatives of many particular Things' are, of course, dispositionally possessed type-ideas, since token-ideas can occur no more than once, and then momentarily. But to say that a type-idea is used representatively is to say that tokens of the type are used representatively. Hence in the elaboration of the theory Locke stressed that the abstract idea before the mind in actual universal thought, in actual conception, is the fully particular token-idea:

the immediate Object of all our Reasoning and Knowledge is nothing but Particulars.... So that the Perception of the Agreement or Disagreement of our particular *Ideas*, is the whole and utmost of all our Knowledge. Universality is but accidental to it, and consists only in this, That the particular *Ideas*, about which it is, are such, as more than one particular Thing can correspond with, and be represented by.⁶⁰

Consequently what is in the primary sense before the mind at different times is not some permanent representative like a mental colour-chart, for there is no such thing on Locke's theory, but distinct tokens of the same precise type. The doctrine is simply that we may pick on any token of the type, whether perceived by the senses or imagined, to be 'the immediate object of our reasoning', and to represent other members of the same sort.⁶¹

Thus to use a particular image as a representative is what it is, on Locke's view, to consider a general attribute. If the favoured particular passes away, we can shift our attention to some other representative exactly resembling the original idea, i.e. to another idea of precisely the same type. We can then perceive, in perceiving its relations to other abstract ideas, the same universal truths as before. Any other men, at any other times, can do the same. It is for that reason, Locke wrote, that the eternal truths are called eternal,

not because they are Eternal Propositions actually formed, and antecedent to the Understanding, that at any time makes them; nor because they are imprinted on the Mind from any patterns, that are any where of them [*sic*] out of the Mind, and existed before: But because being once made, about abstract *Ideas*, so as to be true, they will, whenever they can be supposed to

be made again at any time past or to come, by a Mind having those *Ideas*, always actually be true.⁶²

In other words, universal truths are universal because they are open hypotheticals. In saying so, Locke was echoing the anti-realist tradition.⁶³

Berkeley's famous critique of Locke was therefore based on a misunderstanding or misrepresentation. He took Locke's theory to be that we can have an absolutely distinct and separate image of every distinguishable sensible quality or type of object: as an image of colour or motion without extension, or an image of colour which is not of any determinate colour.⁶⁴ At least one modern attempt to defend Locke has argued that, on a properly sympathetic understanding, abstract ideas are not sensory images at all but objects of a more refined type of thought.⁶⁵ Yet that is in effect to import into our account of Locke's psychology the Cartesian distinction most alien to it, the distinction between imagination and intellect. The truth is that Lockean abstract ideas are images (or rather, perhaps, aspects of images), but are not abstract in quite the way Berkeley supposed.

Locke's concrete language of 'retaining' or 'separating' the abstract idea, of 'leaving out' its concomitants and so forth, might seem to count against this claim. Such language is reminiscent of Gassendi's account of our formation of 'common' ideas, yet even Gassendi, for all the queerness of his general images, glossed 'removed' (*detractis*) with 'or not considered' (*seu non spectatis*). Berkeley made considerable play with a notorious passage, mentioned in another connection in [chapter 6](#) above, in which Locke had been eager to suggest that very abstract thought is out of the reach of children: even the general idea of a triangle is difficult to form, 'for it must be neither Oblique, nor Rectangle, neither Equilateral, Equicrural, nor Scalenon; but all and none of these at once'. This seems to be saying that the general idea has to achieve the distinct standards for Gassendi's 'universal' and 'common' ideas both at once, standards which even Gassendi had said could not be achieved by either separately. But the passage is rhetorical and, as was seen in [chapter 6](#), its point can be otherwise explained. A much more significant passage, reminiscent of the Port Royal *Logic*, points out that logical distinction is possible where full separation, even in thought, is not: 'Many Ideas require others as necessary to their Existence or Conception, which are yet very distinct Ideas. Motion can neither be, nor be conceived without Space.' Furthermore 'a partial consideration is not separating. A Man may consider...Mobility in body without its Extension without thinking of their separation'.⁶⁶ For Locke, then, as for Berkeley himself, abstraction is this kind of 'partial consideration' of the objects of imagination or sensation for the purpose of general thought.

An understanding of abstraction as 'partial consideration' or selective attention is not easy to avoid in the case of certain Lockean abstract ideas. The ideas of unity and existence (which Berkeley called 'the most abstract and incomprehensible of all other')⁶⁷ are said to be suggested by, or brought along

with every idea which comes before the mind. Yet that is to say they cannot be separated from any at all, since whatever was removed, they would still remain. If abstracting were separating, we could continue abstracting unity from existence and existence from unity indefinitely, as if we were dealing with an inexhaustible set of alternately coloured Russian dolls. In fact Locke carefully avoided talk of separation in these cases: the idea of existence is 'suggested' to the mind by ideas and objects which we 'consider' (no doubt partially) 'as actually being there': 'which is, that they exist'. Similarly 'whatever we can consider as one thing, whether a real Being, or *Idea*, suggests' the idea of unity.⁶⁸

The treatment of numbers in the *Essay* is a good example of thought which is more subtle than might be suggested by the figurative language in which it is expressed, and therefore than is generally believed. To have the idea of unity before the mind is to consider something (anything) as one thing. Locke seems to have meant too that to consider the number five is to consider five objects of thought of any kind ('Men, Angels, Actions, Thoughts, everything that doth either exist, or can be imagined'),⁶⁹ but to consider them partially with respect to the point of resemblance between their set or 'collection' and other sets of five: i.e. their number. It is not, that is to say, to consider an independent abstract object or universal. When the number is large, 'names' are necessary, for we have to grasp the number of the 'collection' step by step, counting its members. As he stated explicitly, given the ordered series of numerals all we need to 'perceive' is the difference between next sets, i.e. the difference involved in having one more: '*to reckon right, it is required, 1. That the Mind distinguish carefully two Ideas, which are different one from another only by the addition or subtraction of one Unite. 2. That it retain in Memory the Names*' in their order.⁷⁰ Because the difference between next sets, the addition of a unit is a precise step of which we can have a perfectly clear and distinct idea, and because the series of such steps is matched precisely by the series of numerals, enabling us by counting to grasp the number of a 'great multitude' of individuals which would otherwise be 'a heap in Confusion', number is the subject of a precise demonstrative science.⁷¹ Although Locke sometimes expressed himself in such a concrete way as to seem to confuse numbers with multitudes, there is really nothing in what he says to excite our scorn. Arithmetic is the science of all 'collections' considered merely as such, but it is not his view that numbers are themselves collections of peculiar abstract individuals in the mind called 'units'.

What is mistaken, then, is not the ascription of an image-like or sensory character to Locke's abstract ideas, but the interpretation which takes them to exist in real separation from other ideas. The more we are inclined to view them as non-sensory and fully separate, the more the essential nature of his theory will slip from our grasp. It may not always be helpful to look on Locke as an 'empiricist', but his doctrine that experience supplies the materials of thought does involve a thorough-going and consistent experiential or imagist theory of thought. For Locke the child can 'perceive' a relation between ideas literally with

its senses, knowing before speech ‘the difference between the Ideas of Sweet and Bitter (i.e. that Sweet is not Bitter)’. Explicitly, the child comes to know the truth that three and four makes seven ‘upon the same Grounds, and by the same means, that he knew before, That a Rod and Cherry are not the same thing’, i.e. by perception of a relation between ideas of sensation.⁷² The object of geometrical reasoning may thus literally be a ‘partially considered’ object of sight. It is true that mathematicians’ ‘Demonstrations, which depend on their Ideas, are the same, whether there be any Square or Circle existing in the World, or no’, since they ‘concern not the Existence of any of those Figures’ and so ‘depend not upon sense’, i.e. upon sensitive knowledge of existence. Yet where diagrams are used, they are the particulars ‘perceived’ in reasoning, to its benefit: ‘as soon as the figure is drawn, the Consequences and Demonstration are plain and clear’.⁷³

The doctrine of the *Essay* was thus not so different, with respect to ‘discerning’, from the knowledge-empiricism which Locke had advanced in 1671:

the foundation [of geometry] being all laid in sense viz. sight, the certainty thereof however looked on as the greatest we can or expect to have can be no greater than that of discerning by our eyes, which the very name Demonstration how highly soever magnified for its certainty doth signify.⁷⁴

We can therefore see why it was important for Locke to distinguish sensitive knowledge from intuitive or demonstrative knowledge by its content, as the *perception of particular existence*, rather than by the role played by the senses, as *sense-perception*. For, while it is only in the perception of existence that the senses are for him absolutely necessary (elsewhere imagination can fulfil their role), yet he evidently believed that the senses can be, and are often employed in the perception of other relations between ideas. Thus the theory of abstraction in the *Essay* was above all an attempt to explain the universality, *a priori* cognizability and timelessness of the eternal truths without impugning the sensory character of what is before the mind, and without departing from the general principles of intuitionism. It was in the same tradition as Gassendi’s, but followed Hobbes in not postulating ‘queer’ general ideas. At the same time Locke evidently found Hobbes unconvincing in his claim that language is essential to universal thought.

It is worth considering the internal evidence that Locke was not only aware of the thesis that language is necessary for universal thought, but that he constantly bore it in mind, deliberately and with some reason neutralizing this element in Hobbes’ theory. For Hobbes, those philosophers

err, that say the idea of anything is universal; as if there could be in the mind an image of a man, which were not the image of some one man, but a man simply, which is impossible; for every idea is one, and of one thing.

Consequently 'this word universal is never the name of any thing existent in nature, nor of any idea or phantasm formed in the mind, but always the name of some word or name'.⁷⁵ Yet Hobbes' particular image before the mind in general thought is, in effect, identical with the Lockean representative of a class. Locke had only to see this and to reject the associationist account of reasoning (as he did implicitly in treating 'the association of ideas' as a source of error) in order to conclude that the particular representative idea can as reasonably as its name be called, in virtue of its function, 'universal'. He also, not unreasonably, preferred to believe that our ability to pick out an aspect of a particular in abstraction must be prior to our ability to apply a general name to it on the basis of that aspect or 'consideration'. Unless we could pick out the point of resemblance without the name we could never associate it with the name in the first place. Moreover the principle of unity of the class is not the name, but that point of resemblance in the particulars which is supposed to excite the name. Thus the universality of ideas is prior to the universality of words. As he said in language which suggests a deliberate counterblast to Hobbes, words are general

when used, for Signs of general *Ideas*; and so are applicable indifferently to many particular things; And *Ideas* are general, when they are set up, as the Representatives of many particular Things; but universality belongs not to things themselves, which are all of them particular in their Existence, even those Words, and Ideas, which in their signification, are general.⁷⁶

Correspondingly, while Hobbes' account of truth involved names essentially, Locke elaborated the commonplace distinction between mental and verbal truth, 'as there are two sorts of Signs commonly made use of, *viz. Ideas* and Words'.⁷⁷

It is an interesting and significant aspect of Locke's argument that, having established as he thought the logical priority of ideas and mental predication over words and verbal predication, so that the former are presented as in principle independent of the latter, he then leant over backwards to concede what could remain of the psychological importance of language as an aid to general thought. General truths are 'very seldom apprehended, but as conceived and expressed in Words'.⁷⁸ His chief accounts of abstraction mention 'names' prominently, so that it sometimes seems that the metaphor of a 'standard' appealed just because the abstract idea is a standard for the use of a general word.⁷⁹ Locke suggested too that the purpose of abstraction is to make language more useful by making general names possible, and he found reason to conclude that 'beasts abstract not' from their lack of general words or signs.⁸⁰ His account of number should surely be read as a deliberate concession to a specific nominalist argument trenchantly advanced by Hobbes.⁸¹ Nevertheless the *Essay* states categorically that 'when we make any Propositions within our own Thoughts, about *White* or *Black*, *Sweet* or *Bitter*, a *Triangle* or a *Circle*, we can and often do frame in our Minds the *Ideas* themselves, without reflecting on the Names'.⁸² Thus his view seems to have been that abstraction and general thought can in principle occur

without language, but would probably never have arisen without the need for general words.

Locke was, then, in a theoretically strong position in relation to Hobbes. Yet the doctrines of both contain a fatal flaw recognition of which seems to lie behind later arguments of Hume and Kant. On the approach adopted by Hobbes, Locke and (for that matter) Berkeley, it is the possibility of generalizing the perception of the particular relationship now before the mind to all actual and possible cases relevantly just like the present case which explains the universality, eternity and immutability of the object of universal knowledge. Yet the account only seems to work because it is assumed that we can perceive in the particular case, not only that feature B coexists with feature A, but that B exists because of A. Hobbes' thinker had to observe that the equality of a particular triangle's angles to two particular right angles 'was consequent, not to the length of the sides, nor to any other particular thing in his triangle; but only to this, that the sides were straight, and the angles three'. This consequential relationship must, on the imagist view, be 'observed' or 'perceived' in the particular case and must therefore, presumably, be already present in the particular case, before the act of generalizing to all relevantly similar cases. If it were not already present there would be no way of distinguishing between necessarily related aspects and contingently related aspects of the particular case. That is to say, there would be no way of distinguishing the universalizable perception that this triangle has angles equal to two right angles from the non-universalizable perception that this triangle is red. Consequently the purported explanation of universality in terms of the role of abstract ideas as standards and representatives is a philosophical illusion. It simply glosses over the real difficulty for the imagist, and the nub of the epistemological problem of universality, which in this context is the problem of how sense and imagination could possibly inform us, by an immediate and particular intuition, of *necessity*.

It may be worth noting in passing that, where the universal truth is palpably analytic or, as Locke would have said, 'trifling' or 'verbal', as in 'all triangles have straight sides', it seems possible to explain away the consequential relationship between features as a shadow of the conventional relationship between features and names, i.e. it seems possible to argue, within the terms of theory, that to perceive that the straight-sidedness of a particular figure is consequent upon its triangularity is simply to recognize that the same aspect as excites (or justifies) the name 'straight-sided' helps to excite (or justify) the name 'triangle'. It was presumably for this kind of reason that Hobbes, who wished to see all necessary truths as in some strong sense analytic, felt that names play an essential, more than psychological role in universal thought.⁸³ Yet such an extremely conventionalist model simply fails to fit what Locke, less optimistic than Hobbes about the potentiality of mere analysis, distinguished as 'instructive' truths. A figure's having angles equal to two right angles is a consequence of its triangularity, but the former feature is not a part of what, for

Locke or for Hobbes, conventionally justifies the name 'triangle'. Theorems and axioms leave the imagist-nominalist with a problem, even if definitions do not.

It seems that at least one half of Locke's theory must give way, whether the exclusive status accorded to sense and imagination or the intuitionism, i.e. the doctrine that we can grasp necessary connections by a single perceptual act. Hume discarded the latter, giving imagist theory of abstraction a vigorous twist which seems in general to have gone without clear recognition.

Hume's high praise of Berkeley and his revival of a quasi-Hobbesian nominalism and associationism should not mislead us into supposing that his doctrine is just a version of theirs. He did hold, like Hobbes, that a universal name is associated with impressions and ideas between which there is a point of resemblance. When the language-user hears a universal name or proposition, images of appropriate particulars are excited in his mind. We typically possess a wide range of such images stored up as potentialities, any one of which may be drawn out and become actual 'as we may be prompted by a present design or necessity'. Now when we consider whether a universal sentence expresses a truth, custom brings some of these particular ideas to consciousness. But quite extraordinarily and inexplicably, so Hume wrote, if the sentence happens to be false, just those particular ideas tend to be stimulated which *prove* the sentence false:

Thus shou'd we mention the word, triangle, and form the idea of a particular equilateral one to correspond to it, and shou'd we afterwards assert, *that the three angles of a triangle are equal to each other*, the other individuals of a scalenum and isoceles, which we overlooked at first, immediately crowd upon us, and make us perceive the falshood of this proposition, tho' it be true with relation to that idea, which we had form'd. If the mind suggests not always these ideas upon occasion, it proceeds from some imperfection in its faculties; and such a one as is often the source of false reasoning and sophistry.⁸⁴

Hume had evidently given up the orthodox model for abstraction according to which the mind can extract from any member of a class the material for an all-at-once apprehension of universal truth. He had turned to a radically different solution which is both dispositionalist and falsificationist. There is no such thing as the perception of a necessity or impossibility: there is only the capacity to think up counter-examples. The perception of a relevant relationship in the particular case is quite unproblematic: we perceive that a particular triangle has unequal sides in the same way as we perceive that it is red. Now Hume's routine ascription to 'custom' of the intelligent capacity to think of examples and counter-examples may not be very plausible, but his theory obviously has considerable force. As he himself pointed put, it allows readily for error, avoiding the postulation of an infallible faculty by which universal knowledge is acquired. Moreover, we do seek and employ actual or, more usually, imaginary

counter-examples in *a priori* reasoning, and no general account of such reasoning which neglects that fact has much chance of being right. Yet Hume did nothing to meet the natural assumption that, at least very often when we cannot construct imaginary counter-examples to a given universal proposition, we at the same time 'see why' no counter-example could possibly be constructed. It is the close link which holds together *a priori* knowledge with immediate understanding or insight which drives us (at least in philosophically unprejudiced moments) to use the metaphor of 'seeing' what is necessary or impossible. Hume's doctrine, like more recent dispositionalist theories, does considerably less than justice to the force of that metaphor.

There is another way, however, in which Hume's theory is inadequate, and more demonstrably so. He was no doubt right to reject Locke's disguised assumption that a sensation or image can contain or reveal a necessary connection between particular abstract ideas. Necessity is not an object of sense or Lockean imagination. Yet Hume's own imagist explanation of *a priori* knowledge depends on a similar assumption that sense and imagination can reveal modality. For he assumed, quite explicitly, that whatever can be clearly imagined is *possible*.⁸⁵ Yet the presentation of an object to sense or imagination is no guarantee that an object could exist as presented. Escher's drawings of impossible objects neatly demonstrate the point.⁸⁶ So Hume's move, profoundly interesting and admirable though it is, fails to salvage an imagist theory of our knowledge of the modalities, what is necessary and what is possible.

Abstraction and the ideal of precision

Another criticism has more recently been advanced of an assumption which Locke's theory shares with many others. That is the assumption that the particulars grouped by a general word or concept owe their membership of the class to a precise point of resemblance or feature or aspect equally present in each one of them. It is in virtue of this supposed common aspect that, for Locke, any one of them can represent the others. Yet if we consider the class of red things, it will be seen that they form a continuum within limits such that two red things, A and B, at opposite ends of the continuum, may resemble each other even in respect of their colour less than each resembles other objects falling outside the class. If, for example, A is close to orange and B to mauve, then A will resemble certain orange things more than it resembles certain red things, including B. It follows that the limits of the class cannot be set by 'resemblance'; and indeed similarity, when unspecified, is essentially indefinite and unbounded. Because it draws no line, the assertion that 'A is like B in colour' has no determinate truth-conditions. Unless in a special context, it could be taken only as the expression of a subjective response, an indication of how the speaker is struck which auditors might 'understand', and with which they might sympathize, but which they could not regard as true or false.⁸⁷ The predicative judgement that A and B are both red cannot be reduced to a judgement of similarity: it presupposes a boundary round the class of red things. But it is evident that that boundary is essentially conventional, set by the ordinary meaning of the word 'red'. The same nominalist argument can be extended to many other concepts, although not, as we shall see, to all.

This objection assumes what Locke, if pressed, would have been inclined simply to deny, i.e. that the concept *red* is a well-formed concept. Why that is so will perhaps become clearer in the next section, but Locke was intensely aware, no other philosopher more so, that universal truth and *a priori* knowledge presuppose bounded classes. Yet he argued explicitly that public language is the last thing which can be supposed to determine boundaries which are sufficiently definite and unambiguous.⁸⁸ He embraced, unreluctantly enough, what might seem the only alternative: the class represented by the particular abstract idea is determined by *precise* resemblance to the idea. Only in that way could he

suppose that particulars have within them what is required for the intuition of universal truth.

The paradigm was mathematical: particular triangles resemble one another precisely, not of course in all respects but ‘partially considered’, in the particular respect which determines that they are all triangles. Each one of them *qua* triangle has just what every other one has, and a problem over the boundaries of the class of triangles does not arise for his theory as it does over the boundaries of the class of red things. If we say that all red things have redness in common, ‘redness’ is not a feature which is precisely the same in each example: it is not a point of precise resemblance. Locke’s response to such considerations was to uphold the mathematical paradigm incorporated into his programme for the reform of scientific language. The ideal of precision was thus, like intuitionism, one of the fundamental features of his thought with which the doctrine of abstraction achieved a neat and fully self-conscious fit:

Now because we *cannot be certain of the Truth of any general Proposition, unless we know the precise bounds and extent of the Species its Terms stand for*, it is necessary we should know the Essence of each *Species*, which is that which constitutes and bounds it. This, in all simple *Ideas* and Modes, is not hard to do. For... the abstract *Idea*, which the general Term stands for, being the sole Essence and Boundary, that is or can be supposed, of the *Species*, there can be no doubt, how far the *Species* extends, or what Things are comprehended under each Term: which, ’tis evident, are all, that have an exact conformity with the *Idea* it stands for, and no other.⁸⁹

Nevertheless, Locke did not elsewhere escape from a position of unresolved ambivalence with respect to ideas of colours—which were, after all, his paradigm simple ideas. He sometimes said what doctrine required: for example, that words like ‘red’ do not, strictly speaking, name simple ideas but rather bundle up, by a loose convention, a large number of distinct shades, the true simple ideas.⁹⁰ That would seem to concede a kind of nominalist explanation of ordinary, sloppy usage, while upholding the more precise abstractionist ideal. Yet he also invoked the notion of a ‘simple mode’, allowing that shades of red ‘are considered but as different degrees of the same simple *Idea*’.⁹¹ The point would seem important, since ideas of simple modes, unlike simple ideas, can be constructed. Certainly Locke seemed to take seriously the notion of degrees of a simple idea, for from a simple idea, such as *white*, we are supposed able to form a relative idea, such as *whiter*.⁹²

Another problem which Locke seems to have recognized is that shades of colour are not open to precise discrimination each from the next. A similar difficulty arose for him in respect of ‘distinct *Ideas* of every the least excess in Extension’, but that he took to be resolved by the possibility of geometrical proof of perfect equality.⁹³ The problem of finding a range of distinct, precisely

identifiable shades he left unresolved: yet in another connection Locke did not deny, but stressed, the 'distinctness' of simple ideas.⁹⁴ Generally he simply ignored the admitted failure of psychological reality to come up to the logical ideal. Yet the truth seems to be that this failure is not merely psychological, but that the ideal itself is inappropriate. For the difference between 'precise' mathematical concepts and 'imprecise' concepts does not stem from any variation in our powers of sensory discrimination.

The application of number to reality comes, logically speaking, after the discrimination or individuation of things to be counted. Any difficulty in deciding how many things there are will therefore be attributable to the difficulty in discriminating things. It is for that reason that number concepts themselves seem serenely precise and unsubject to boundary problems. The application of geometrical concepts to the world is not in the same way second order, but it is peculiarly tentative or approximate or 'for all intents and purposes'. Any observation or measurement which shows that two angles are not precisely equal, or that a figure is not precisely square, shows that the angles are not equal, the figure not square. But then we do not expect any actual pair of angles to be more than approximately equal, or a figure to be better than approximately square. For the purpose of applying our geometrical calculation we simply postulate perfect equality, perfect squareness or whatever, and are unsurprised when reality only approximates to our conclusion too. Otherwise we are prepared to call something square when for the purpose in question the imperfection can be ignored. It does not follow that squareness is relative or a matter of degree. Still less does it follow that, as it is often supposed, there are two concepts of squareness, a precise mathematical one and a vague empirical one equivalent to 'roughly square'. Squareness is, on the contrary, always a precise ideal to which things may approximate, but to which they are not in general expected to do more than approximate.

The assumption that precision is always an appropriate ideal underlay a large part of Locke's programme for the advancement of science to its practicable limits. He did not accept that there is any impediment in principle to our constructing a language in which all types of terms have the precision of mathematics, i.e. a language such that, because all ideas expressed in it are precisely determined, all universal propositions expressed are in principle capable of being 'perceived' to be true or false, and so of being the material of a demonstrative science. Locke did of course believe that, largely because of the limitations of our senses, a 'science' of substances is a practical impossibility. Yet that barrier was not supposed to concern modes or relations, or (except avoidably) to relate to precision. One of the importantly novel themes of twentieth-century philosophy has arisen from the claim, in effect, that Locke was here wrong. First, it is arguable that words like 'red' or, to take Wittgenstein's famous example, 'game', are necessarily typical of those terms in a natural language which Locke classed as names of simple ideas, mixed modes and relations. It seems that mathematical precision is necessarily special, and the

possibility of mathematical or quasi-mathematical *a priori* science is to be explained (as was very schematically attempted in the previous paragraph for arithmetic and geometry) as a peculiarity of special areas of language, of special symbolic techniques. Second, to anticipate later discussion, names of substances such as 'horse' or 'gold' are semantically so different from either 'red' or 'square' that their precision, if that means the precise bounds of their extension, depends as much or more on nature than on us.

It is interesting that Berkeley recognized the role of the ideal of precision in the doctrine which in other respects he misconstrued. In the Draft of the Introduction to the *Principles*, he argued that a general name is made the sign of many particulars 'between which there is some likeness', but he denied that the class need have 'any precise bounds or limits at all': 'for if they had I do not see, how there could be those doubts and scruples, about the sorting of particular things'. Precise boundaries are unnecessary, 'language being made by and for the common use of men, who do not ordinarily take notice of the minuter and less considerable differences of things.' Elsewhere, even more suggestively, he implied that there need be no single defining feature present in each and every member of the class, so that the question arises for the doctrine of abstraction whether there is 'in this your abstract idea of man, the idea of eyes, or ears, or nose, or legs, or arms': 'there being particular men that want, some arms, some legs, some noses etc.'. Evidently Locke must accept 'an odd and frightful figure, the idea of a man without all these'.⁹⁵

Yet Berkeley was unable to make use of the intuitions behind these arguments. In the published work they were omitted, and attention was focused on the traditional, supposedly crucial example of geometrical proof that had been discussed by Hobbes, whose penchant for precision was hardly exceeded by Locke's. Moreover, Berkeley never stated that mathematical precision is inappropriate even as an ideal for many concepts. Philosophy had to wait for Wittgenstein before that task was clearly done. After Wittgenstein, however, the fact that the boundaries embodied in language are often imprecise, indefinite and open to dispute need not be taken as an argument that they ought to be replaced, or even that they could be replaced, by something better which is not dependent on the common tongue. Yet, as I have suggested, that is not to say that all class boundaries marked by language are dependent on language. The class of men, Berkeley's example, has, as a natural species, a boundary supplied by nature rather than by language. Still less can we conclude that any realist theory of universals is false. The problem of universals is not one to be approached with sweeping theories in hand, whether logical, epistemological or ontological. For, as it will be argued in Volume II, we need to arrive at a satisfactory theory of categories before we can arrive at a satisfactory explanation of the logical, epistemological and ontological status of this or that category of 'universals'.

Intuition and innate knowledge

How should we assess Locke's intuitionism, and the analogy of 'seeing' or 'perceiving' necessary connections? From the general discussion of knowledge in [chapter 15](#), above, it should already be clear that what is chiefly at issue is the 'perspicuity' of necessary knowledge. Is there a subjective condition of the apprehension of necessary truth at least partly analogous to sensation, a perspicuous 'appearance' of necessity which may in particular cases be either accepted or rejected, but which is in general veridical and which naturally and normally issues in uninferred belief?

A relevant question here is whether there is room for a reasonably strong notion of *a priori* illusions, as opposed to mere slips, mistakes and false beliefs. At first blush there do seem to be fairly close analogues even to the standard perceptual illusions. Few would assent to the paradoxical conclusion of Zeno that motion is impossible, but at the same time his arguments seem, in a sense, convincing. It may indeed be that no one has successfully explained what is wrong with them. Even if we can clearly identify a false premise or fallacy in such an argument, the argument can retain its appearance of cogency rather as an optical illusion may remain even after we know it for what it is. That appearance cannot be explained as a mere inclination to hold a belief, now repressed, any more than a sensory appearance could be so explained. For in *a priori* 'appearance' we are inclined to believe the proposition which appears necessary in some sense *because* of that appearance, in that our belief does not arise subjectively as a mere hunch, true or false. It is because I 'see why', *if A is larger than B, and B than C, then A is larger than C* that I am inclined to believe that proposition. And, as Locke allowed in his account of 'assent' but failed to recognize with respect to 'perception', I can stand back even from the most striking appearance of cogency in order to reflect critically upon it. Like the deliverances of sense, *a priori* evidence is defeasible, and the evidence of a necessary truth may not be subjectively different in kind from the 'evidence' of the premises and steps of some fallacious argument which I find equally intelligible and convincing. Compare arguments on opposite sides of any unresolved philosophical controversy.

It is a popular view that *a priori* 'evidence' can be nothing more than a strong 'feeling of conviction' or 'feeling of understanding' which accompanies belief. If that were true, there would be nothing to stop any rigmarole from appearing evident or necessary to us. Yet if someone in some sense 'felt' that the sentence 'If A is a table, then B is a chair', or 'A smell of cheese prevails throughout', is universally true, necessary and intelligible, he would not thereby be suffering from an illusion of reason. He would have to be deluded or the victim of mental confusion, since for an illusion of reason to occur the proposition must *appear* necessary. A mere piece of nonsense, or for that matter some palpably contingent proposition, could not do that. For there to be an *appearance* of necessity there must be something in the proposition or argument considered which misleads. That means that at some level the proposition or argument must be understood. Some sort of intelligent understanding is a condition of this sort of misunderstanding, rather as we must perceive what we misperceive.

Yet here the limitations of the analogy with sense-perception begin to show themselves. A sensory hallucination is at least in principle possible which is utterly falsidical and unrelated to any external reality, a mere figment of the brain, but which is in itself subjectively exactly the same as a veridical sense-experience. There is no analogue to such a possible phenomenon in the case of *a priori* knowledge and belief. There is 'appearance', but the appearance of necessity is not, like sensory appearance, the end product of a discrete, 'modular'⁹⁶ mechanism of its very nature preliminary to judgement, assessment and belief. It is the product, as it seems, of just the same faculty as must arrive at a final judgement. Although the appearance is more than a mere inclination to believe, in that it involves some awareness of why such an inclination exists, it is not, like sensation, a presentation of a state of affairs upon which belief may be grounded.

Like all good analogies, then, the analogy between *a priori* understanding and sense-perception is less than complete, but it is not therefore to be dismissed. Just as perception and perceptual knowledge necessarily involve the operation of a cognitive faculty and mechanism, so *a priori* understanding necessarily involves the operation of a cognitive faculty and mechanism. Let us call the latter faculty, without prejudice, 'intelligence'. As the standard sensory illusions involve the mechanisms of sense, so illusions of reason involve the mechanism, whatever it may be, of intelligence. Indeed, it may require considerable intelligence in order to follow some sophistical argument sufficiently well to be deceived by it. As perceptual knowledge is 'perspicuous', so is *a priori* knowledge. In talking of the perspicuity or 'evidence' of *a priori* knowledge, we are talking of the difference between intelligent insight and some reliable hunch-producing mechanism. We are talking of a condition of 'primary' rather than 'secondary' *a priori* knowledge, in the sense of these terms explained in [chapter 15](#), above.

Advocates of the 'causal theory of knowledge' have sometimes shied away from the task of identifying a causal condition of *a priori* knowledge just

because they have seen the causal condition of factual knowledge as a matter of the right causal route from the fact known to the belief, a formulation inapplicable to *a priori* knowledge. *A priori* facts have no natural effects. Another consequence of their approach has been the entangling of the causal condition of knowledge with that other objective condition of knowledge, the truth of what is believed. For if the causal condition consists in a relation between fact and belief, it would seem to entail or require that the belief be true. If, however, we keep separate the causal factors which the causal theory tends to lump together (the operations of cognitive faculties, the relations between effects inferred from causes or causes inferred from effects, the passage of information from one person to another and so forth), it becomes possible to find a causal condition of *a priori* knowledge: roughly, that it should be the product of intelligence. It also becomes possible to see that the causal condition of knowledge may sometimes be fulfilled when the condition of truth is unfulfilled. For example, since the standard sensory illusions arise when the senses are functioning normally and well (even if they fall short by comparison with some ideal mechanism) it can be said that, when they occur, the causal condition of knowledge is satisfied even though the condition of truth is not. Analogously, the reason why belief in a true *a priori* conclusion arrived at by a fallacious (but ingenious) argument or from a false (but intellectually plausible) premise does not constitute knowledge is not because the causality of the belief is not in general of the right kind, but simply because of the falsity of the premise or the fallacy in the argument. There need be nothing wrong with the intelligence of someone who accepts a sophistical *a priori* argument; otherwise most philosophers would suffer from defective intelligence. Indeed, only by separating out truth and validity on the one hand from appropriate causality on the other can we avoid defining rational mistakes out of existence and so plunging back into the Cartesian psychology, shared in this respect by Locke, which attributes error solely to such irrational factors as passion and prejudice. That is to revert to the sufficiently discredited notion that reason is in itself infallible, a natural faculty defined by a logical criterion. The more comfortable view is that intelligent mistakes are possible, and that the deliverances of intelligence, like the deliverances of sense, are defeasible.

Despite that difference from Locke's conception of 'evidence', the notion here proposed of the perspicuity of *a priori* intuition is perhaps sufficiently like his to encourage sympathy for his attacks on the doctrine of innate general principles. The core of his case was this: not only is it quite superfluous to suppose that any of the truths that we know by intuition or demonstration were already innate in us, but the theory that we knew a truth innately before it came to consciousness and was in the full sense perceived is self-contradictory.⁹⁷ Ever since his ink was first dry, he has been accused of refuting only an absurd version of the theory of innate knowledge, a version according to which the knowledge is from the first actual or explicit rather than dispositional or, as Descartes himself had put it in response to objection, innate only in the sense in which we say that in certain

families 'certain diseases, such as gout and the stone, are innate'.⁹⁸ Yet this criticism of Locke is entirely misconceived. It is true that in the somewhat knock-about rhetoric of Book I he derived some entertainment at the expense of a straw figure, a non-existent philosopher who, believing the maxims of logic innate, 'will say, Children join in these general abstract Speculations with their sucking Bottles, and their Rattles', and who fails to appreciate that 'such kind of general Propositions, are seldom mentioned in the Huts of Indians'.⁹⁹ Yet the theme of the dialectical argument in which these passages occur constitutes a direct attack on the dispositional theory:

if these Words (*to be in the Understanding*) have any Propriety, they signify to be understood. So that, to be in the Understanding, and, not to be understood; to be in the Mind, and never to be perceived, is all one, as to say, any thing is, and is not, in the Mind or Understanding.¹⁰⁰

No one, of course, held a 'sucking-bottle' version of the theory, and Locke's point was that only such an obviously false theory makes sense. For we do not know a general truth unless we perceive it, i.e. consciously understand why it is true. That there are truths which, by a certain age, we readily or immediately understand when they are proposed to us supplies no reason to postulate anything innate but a general capacity for rational understanding: 'Universal and ready assent, upon hearing and understanding the Terms, is (I grant) a mark of self-evidence: but self-evidence [depends] not on innate Impressions, but on something else (as we shall show hereafter)'.¹⁰¹

Leibniz, responding to Locke, offered the distinctly patronizing, standardly Platonist reminder that there can be no difficulty in the notion of dispositional knowledge, since what is in the memory is undeniably known dispositionally.¹⁰² Yet that is no shield against the thrust of Locke's argument: if an infant or anyone else understands something, then, first, that understanding must be due to the operation of a cognitive faculty rather than a mere disposition to believe; and, secondly (and *ipso facto*), it must either be or have been conscious and explicit. Some of the clearer anti-dispositionalist comments in Book I were added to the second edition, but even in the first edition Locke had explained the difference between 'actual' and 'habitual' knowledge in terms which rule out innate dispositional knowledge. Habitual knowledge is a man's knowledge 'of all those Truths, which are lodg'd in his Memory by a foregoing clear and full perception'.¹⁰³ But in the concept of innate knowledge the connection between knowledge and 'evidence' or 'perception', the conscious functioning of our cognitive faculties, is broken completely, in effect if not in intention. Beliefs resulting from innate tendencies, however reasonably attributable to God's goodness (and therefore trustworthy) they might be, could never (even if we could know that such tendencies existed) have satisfied Locke's conception of reason or the light of nature.¹⁰⁴ Hence to offer to explain our knowledge of certain principles as being innate is precisely to fail to explain what makes it

knowledge. One might say with little paradox that Locke rejected innatism, not because it is part and parcel of 'rationalism', but because it is insufficiently rationalist. At any rate, he found it no less opposed to a consistently intuitionist view of reason than to that other plank in his anti-Cartesian platform, his reasonable conviction that the senses play a more fundamental and yet more limiting epistemological role than Descartes had allowed.

If the argument of this section is correct, then a qualified form of intuitionism is true. Yet there is still need of an explanation of what *a priori* intuition is. If it is an exercise of some form of reason or intelligence, then in what context is the intelligence exercised and towards what objects is it directed? In order to answer these questions it is necessary to turn to the topic of language and meaning.

Locke on meaning and some modern criticisms

Locke's account of the function of language in communication is much like Hobbes':

Men learn Names, and use them in Talk with others, only that they may be understood: which is then done, when by Use or Consent, the Sound I make by the Organs of Speech, excites in another Man's Mind, who hears it, the *Idea* I apply it to in mine, when I speak it.¹⁰⁵

The model may seem to be no more than a corollary of the traditional doctrine of concepts and terms, yet Locke saw it as a potentially radical thesis, a weapon to be employed against two traditional errors. The first is the assumption that the words in a language mean the same for all those who speak it: the assumption of those who

think it enough, that they use a Word, as they imagine, in the common Acceptation of that Language; in which case they suppose, that the *Idea*, they make it a Sign of, is precisely the same, to which the Understanding Men of that Country apply that Name.¹⁰⁶

The second error is the assumption which many make that they talk of 'Things as really they are': i.e. that, as well as standing for ideas, 'their Words stand also for the reality of Things'.¹⁰⁷ By this Locke meant the assumption that words get their meaning directly by naming things as they are in themselves, unmediated by ideas, or things as we conceive of them.

These two mistakes about language, according to the argument of Book III, are common causes of everyday confusion, but they are also associated with certain false philosophical doctrines. The first mistake Locke perhaps linked particularly with the doctrine that there are common moral notions embodied in language: while the second, he thought, is epitomized in the Aristotelian doctrine that there are substances out there, distinguished by their essences into species and genera capable of being identified and named by us. Against these two assumptions Locke insisted that names 'in every Man's Mouth, stand for the

ideas he has, and which he would express by them'. The names anyone uses must be ascribed to 'the Ideas that he has, and not to Ideas that he has not'. Yet the philosophical views Locke attacked were not rejected by him solely for the reason that they contravene that principle. Although he regarded it as hardly more than common sense that words have the meaning given to them by the speaker, he recognized that appeal to that principle required more specific support in the case of 'the Names of mixed Modes, and substances, in particular'.¹⁰⁸ In effect the account of meaning was just one element in an elaborately constructed anti-Platonic, anti-Aristotelian argument. It was not, that is to say, the artless and innocent assumption which many today take it to have been, but one component of a complex, overarching thesis which incorporated his intuitionism, his particular sense of the need for clarity and precision in scientific and ethical discourse, and his highly theoretical view of the problems of natural classification.

It is interesting and, indeed, remarkable that the two assumptions about meaning which Locke picked out as sources of confusion and philosophical error correspond closely to two of the most effective arguments to have been brought in recent years against the general model which he advocated. According to one of these arguments, vigorously propounded by Hilary Putnam in particular,¹⁰⁹ the meaning of what we say is independent of our inner states to the extent that a language embodies the knowledge and life of the whole community which speaks it. 'Nuclear fusion', 'entailment', 'molybdenum', 'tiger' have meanings which no expression could have in a linguistic community lacking an appropriate level of knowledge. 'Matriculation', 'engagement', 'checkmate', 'sheep-dipping', 'ostracism' owe theirs to practices within the linguistic community, or at least to knowledge of practices in other communities. The terms of dead theories and doctrine belong to the same general class of words. Yet not every speaker of the language has the knowledge, or understands the theories or practices, which give these expressions their meaning. In the case of very many expressions the experts are vastly outnumbered by the laymen. The question therefore arises how much knowledge the individual speaker must have if he can properly be said to understand, and to use with understanding, such an expression in the common language. This question can be sharpened if we consider that assertion and the acceptance of an assertion are impossible without meaning and understanding. How much knowledge of physics, then, is a necessary condition of my *asserting* that a certain power station employs nuclear fission rather than fusion, or of my *believing* that it does if I am told so?

It seems reasonable to allow that such an assertion and such a belief are not confined to the expert, but are within the power of the moderately informed layman. That answer conforms to the natural and common-sense interpretation of a test of linguistic capacity. Someone who thinks that an onyx is an antelope will fail that question in a vocabulary test, but if they know that onyx is a mineral it seems reasonable that they should pass, while to know that it is a semi-precious stone is surely enough. It would be too much to require knowledge of its

chemical composition, or the capacity to distinguish onyx from other stones. It follows that someone may be held to understand each of two words with different meanings, e.g. the names of the chess pieces, 'pawn' and 'rook', without knowing in what ways they differ. It is within their power to assert falsely that one object is a pawn and to believe truly that another is a rook without knowing the difference between a rook and a pawn, so long, at least, as they know that these things are different. All that is necessary is to have a certain awareness of chess together with knowledge that these are the names of different pieces in the game.

On such a view of language, 'meaning' is a social phenomenon, dependent on what Putnam calls the 'division of linguistic labour' in the community. The public language embodies the public stock of knowledge; and what a word means in the layman's mouth (provided that he has sufficient conception of its role to mean anything) is the public meaning determined by the expert. Thus the layman may have some parasitic share in the knowledge of the expert. He may properly be said to know a number of things about nuclear fusion, for example, without having expert or 'real' knowledge of what nuclear fusion is. If all this is correct, then Locke's model for meaning would seem to be seriously wrong. What primarily determines the meaning of what we say and the content of expressed belief will not be our 'ideas', taken as distinct mental concomitants (whether occurrent or dispositional) of our words, but the conventional meaning of our words in the public language we speak.

Perhaps the first thing to be said in defence of Locke's model is that its purpose was not descriptive but normative. It was designed to portray only an ideal case of *real* or *full* communication involving *truly* meaningful speech and genuine understanding. It was Locke's contention that such cases are relatively rare, and that ordinary conversation takes place all the time which at best approximates to the ideal, and often falls deplorably short of it. In effect, he embraced the paradox of denying the existence of a common tongue except as a common set of words to which different individuals annex different meanings. There are two classes of predicate which he significantly excepted from the general babel: names of simple ideas and names of simple modes. The former can only be learned ostensively, while the latter comprise the precise terms of the paradigm *a priori* sciences. Hence nearly everyone who is not blind has an agreed and determined idea of yellow, while no one 'that had a mind to understand them, mistook the ordinary meaning of Seven, or a *triangle*'. On the other hand, 'there are few names of complex *Ideas*, which any two Men use for the same just precise Collection'.¹¹⁰ The context of this claim was a programme of linguistic reform, not original to Locke but characteristic of seventeenth-century thought, the specific nature and pretensions of which are reflected in his adaptation of the Cartesian expression 'clear and distinct'.

For Descartes, as we have seen, the expression connoted a sort of simple perspicuity or intelligibility possessed by the object of the understanding. The Port Royal *Logic* primarily followed Descartes' usage, but took note of a second sort of confusion, the confusion which is consequent upon unsteady relationships

between ideas and words.¹¹¹ Locke, no doubt influenced by the *Logic*, also took confusion and distinctness in two senses, but for him it is the second which is the more important. With respect to the first sense, whereas the perspicuity of ideas for Descartes and Arnauld is in relation to the intellect, so that clarity and distinctness (the former defined in terms of the latter)¹¹² are functions of scientific understanding, for Locke the perspicuity of ideas is simply sensory or phenomenal, quite separate from understanding and propositional thought. 'Clarity' he defined as the property possessed by simple ideas when they are fit representatives of the quality in the object: 'such as the Objects themselves, from whence they were taken, did or might, in a well-ordered sensation present them'. Complex ideas are clear when they are composed of clear simple ideas in a 'determinate and certain' way. 'Distinctness', understood in this first phenomenal sense, is a property vacuously possessed by all ideas whatsoever, since no idea can seem to be other than it is.¹¹³

The Cartesian notion having been summarily rejected and replaced from the point of view of imagism, Locke turned to the other, theoretically more interesting sense of 'distinct'. We have 'distinct' ideas in so far as the ideas which we associate with different words supposed to have different meanings are themselves different, and so capable of distinguishing different objects. To use Locke's example, if a man takes a 'leopard' to be a different animal from a lynx, but the only idea he has of either is of a beast (or, no doubt, a cat) with spots, then his ideas of them are indistinguishable and so 'confused'. It is another source of confusion if someone associates different ideas with the same word at different times. This private confusion, an improper or disorderly relation between words and ideas, stands in the way of the individual's knowledge: '*Ideas*, as ranked under Names, being those that for the most part Men reason of within themselves'. It leads on to public confusion, for of necessity verbally designated ideas 'are always those which they commune about, with others'.¹¹⁴

To understand Locke's ideal model for communication we must recognize that for him it ought characteristically to be the communication of *knowledge*, dependent on community of the materials of knowledge and belief. According to a striking metaphor several times repeated, speech is 'the common Conduit, whereby the Improvements of knowledge are conveyed from one Man, and one generation to another'.¹¹⁵ Without common ideas, Locke assumed, common knowledge is impossible. Beside his contention that it is impossible to teach someone blind from birth the meaning of the name of such a simple idea as blue, we can set the not implausible suggestion that it is impossible to tell such a person that the sky is blue, i.e. to convey to them what we know when we know that the sky is blue. If, in the light of Putnam's doctrine of the 'division of linguistic labour', we allow that knowledge may be conveyed in such a case, as from expert to layman, then it must be admitted that it is knowledge oddly impoverished in transit: not just in its slipping, in Locke's senses of the terms, from 'knowledge' to 'belief', but in its content. Moreover, the blind man cannot, as we can, see the point of certain general statements about colours. In an

important sense he cannot see *why* what is white is necessarily not black, or why it is true that scarlet is more vivid than brown, or is 'like the sound of a trumpet'.¹¹⁶ For Locke, that is simply because such a man cannot *understand* the terms of those propositions, and so cannot even *believe* them.

Locke's conception of general knowledge was, as we have seen, tied to the mathematical ideal. If we adopt the thesis of the 'division of linguistic labour', then it would seem to follow that someone who knows only that a 'rhombus' is a geometrical figure, and that 'parallel' is a geometrical term, might nevertheless be supposed to understand the lemmas of a proof containing these terms, and might believe them to be true, even though he is not in a position to understand or follow the proof. Locke's view, however, which is here significantly at its most plausible, was that such an auditor is told nothing, and that his knowledge, unless of words, is not increased: 'he, that has not a perfect, and clear *Idea* of those Angles, or Figures of which he desires to know any thing, is utterly thereby incapable of any Knowledge about them'.¹¹⁷ Locke's understanding hearer must have the materials, at least, of rational intuition and demonstration, of *a priori* understanding. Otherwise 'the great Conduit, whereby Men convey their Discoveries, Reasonings, and Knowledge, from one to another' has failed in the case.

We can, then, chart the continual influence on Locke's theory of meaning both of his paradigms for 'simplicity' and of his mathematical ideal of precision and understanding. But we can also chart his sensitivity to the opening for a notion of public meaning which implies that linguistic competence is a skill essentially defined by the practice, and parasitic on the knowledge, of the community as a whole. Of course the notion is one which he wholeheartedly rejected, but he gave it a kind of recognition in, for example, his elaborate distinction between the 'civil' and the 'philosophical' (i.e. scientific) use of words. The former is 'such a communication of Thoughts and *Ideas* by Words as may serve for the upholding common Conversation and Commerce, about the ordinary Affairs and Conveniencies of civil Life', whereas the latter is such a use of words 'as may serve to convey the precise Notions of Things, and to express, in general Propositions, certain and undoubted Truths, which the Mind may rest upon, and be satisfied with, in its search after true knowledge'.¹¹⁸ Locke's programme of reform had to do only with the philosophical use of language, and his ideal for communication arose out of his proposals for the improvement of language as the medium of reason, i.e. of systematic scientific and ethical thought. It would be ridiculous, he wrote, 'to attempt the perfect *Reforming* the Languages of the World', or to require that 'all Men should have the same Notions, and should talk of nothing but what they have clear and distinct *Ideas* of'. For that, men would have 'to be very knowing, or very silent'.¹¹⁹

Nevertheless it would obviously be going too far to ascribe to Locke two distinct conceptions of linguistic competence. He did recognize that in ordinary life it is normal to pick up words, taking them over from others without what he would have counted as 'clear and distinct ideas'. Yet he deprecated the habit:

men ‘begin at the wrong end, learning Words first, and perfectly’, ‘contenting themselves with the same Words other People use’.¹²⁰ He could not find in this procedure an acceptable paradigm of what it is to learn a common tongue. Its consequence he always regarded as a kind of parrot speech which in its worst manifestations leads people even to fight ‘for those Opinions, they never were convinced of...no, nor ever had so much as floating in their Heads’.¹²¹ Mere words are not enough for belief and knowledge. Indeed, it might almost be said that Locke’s approach to language sprang out of his experience of all kinds of ignorant bigotry, through his passionate rejection of the possibility of knowledge that is merely second-hand: ‘we may as rationally hope to see with other Mens Eyes, as to know by other Mens Understandings’. To endeavour to take over the knowledge and language of purported experts without their understanding is servile and useless reverence for authority: ‘Such borrowed Wealth, like Fairy-money, though it were Gold in the hand from which he received it, will be but Leaves and Dust when it comes to use’.¹²² It is after all hardly to be expected that a leading proponent of the Enlightenment should have shared the deep, unswerving faith of its present-day beneficiaries in knowledge as a social achievement. Which is not to say that he did not advocate co-operation. That was the point of linguistic reform.

To put it roughly, one approach to language assumes that communication is normal, so that the conditions of understanding and of linguistic competence, the elusive subject-matter of semantics, should be sought in the conditions of ordinary conversation, a social activity performed against the background of a public store of knowledge, theory and belief. Locke’s interesting and certainly influential alternative was to see a natural language as an essentially indeterminate, adaptable tool, only capable of conveying exact and perspicuous information if special measures are taken to achieve precise meaning, and shared understanding of that meaning. The burden is continually on the individual speaker to explain precisely what is going on, and on the hearer to require such explanation wherever necessary.¹²³ ‘Common use’ may be a necessary basis for communication to get started,¹²⁴ but reference to a public rule is ‘a very variable standard’.¹²⁵ In reality, Locke did not believe in a single public standard at all, but in a host of overlapping private uses: ‘in this respect, the Knowing, and the Ignorant; the Learned, and Unlearned, use the *Words* they speak (with any meaning) all alike. They, *in every Man’s Mouth, stand for the Ideas he has, and which he would express by them.*’ Thus, if they were asked, people would define ‘gold’ variously according to their different knowledge of it. The name is the same, but ‘it is evident, that each can apply it only to his own *Idea*’.¹²⁶

Such extreme semantic individualism has recently been undermined, not only by the notion of a ‘division of linguistic labour’, but also by an argument for what is in effect the second assumption about names excoriated by Locke: the assumption that the meaning of such a word as ‘gold’ is determined directly by its relationship with the ‘natural kind’ of which it is the name, rather than by any mental conception or idea which we may have of that kind of thing. On the view

opposed to Locke's, 'gold' has a unitary common use or meaning in a community provided that all members of the community use it as the general name of the same kind of stuff. In this respect, as Saul Kripke in particular has argued, it is comparable with a proper name. Provided that the old woman who knew Tom Brown only as a child, and the child who knew him seventy years later only as an old man, both used the name 'Tom Brown' for the same individual, then they both used it in the same way and with the same meaning. That is so despite the fact that their 'conceptions' of Tom Brown, and their answers to the question 'Who is Tom Brown?' would naturally have taken very different forms. So people may have very different knowledge of gold and employ very different criteria for distinguishing gold from other substances, and yet, just because they all identify in their different ways the same chemical substance, all use its names, 'gold', 'gelt', 'l'or' etc., with the one common meaning. This line of thought, also pursued by Putnam, combines neatly with the thesis of a division of linguistic labour. If there are experts who can identify, by whatever means, distinct substances which they call 'porphyry' and 'agate', then a layman can derivatively use these names with the same two public meanings, even if he has mistaken, not to say indistinct and confused conceptions of the substances of which they are the names. If both lines of thought are correct, then Locke's notion of the 'idea' as the necessary intermediary, for each individual speaker, between word and object can be excluded from any account of the meaning of words such as 'gold'.

Some apology will be given at a more appropriate place in Volume II for Locke's denial that the names of substances stand for natural species. Once again his position rested in part on an ideal of language as the conduit of knowledge. He was objecting to the traditional assumption that 'gold', 'horse' and the like, having been introduced on the basis of perceived resemblances and differences, are fit to serve as the terms of an explanatory, intuitively evident science. Just as we cannot reason and have scientific knowledge about a pentagon without having a determinate idea of what a pentagon is essentially (simply to have picked up the word 'pentagon', while aware that it is the name of a figure, will not do), so we cannot reason scientifically about gold without having a determinate idea of the essence of gold. Yet since, in Locke's view, no one has (or then had) such an idea of gold (and *a fortiori* not everyone has it), all that can figure in anyone's reasoning and knowledge is the inadequate and arbitrary experiential idea of gold which that person has. Locke's semantic individualism hangs together with his rejection of the view that we can introduce terms into the language, or explain existing terms, simply by naming natural kinds. And, in the end, both may be traced to his individualistic intuitionism: crudely, to his view of knowing as seeing for ourselves. No critique of his account of meaning will be adequate, or even strictly to the point, which does not get to grips with the question of the nature of *a priori* knowledge, and the role of language in relation to *a priori* knowledge.

Reflections on understanding and imagination

The problem of *a priori* knowledge calls for an account of the interpenetration of rationality or intelligence and linguistic competence. As a first contribution to such an account something should be said on the nature and phenomenology of understanding in general, and of understanding speech in particular. These issues can be approached through a familiar line of objection to Locke's model for meaning which has not yet been mentioned.

It is often said that if it were true that all names 'properly and immediately signify nothing but the *Ideas*, that are in the Mind of the Speaker',¹²⁷ we could never know what other people meant or whether they understood us, just because the 'ideas' in their minds would be inaccessible to us. A further argument stresses performance as the criterion of meaning and understanding. Whatever may be supposed to take place in the speaker's or hearer's 'mind' or consciousness, we actually tell whether the one speaks and the other hears with understanding from the context of their behaviour—from the appropriateness of the hearer's response, from the evidence of their general linguistic capacity which is provided by their utterances and responses over time, and so forth. The conclusion of such an argument is the denial that any 'mental' or 'inner' process or event whatsoever is either a necessary or a sufficient condition of understanding. Modifications of consciousness, whether pictorial images, verbal images, 'feelings' of understanding or whatever are at most, in Wittgenstein's words, 'more or less characteristic accompaniments or manifestations of understanding'.¹²⁸ Roughly, understanding a particular speech is held to consist in a capacity or tendency to perform appropriate verbal or non-verbal actions. A somewhat bewildering feature of these arguments is that they are often advanced by philosophers apparently uncommitted to a blanket behaviourism or behaviouristic functionalism. The existence of modifications of consciousness, e.g. sensations and images, is not denied, but only their essential relevance to meaning and understanding. Yet it is difficult to see how the arguments, if they were at all cogent, would do less than prove general behaviourism true. Even the first argument would seem capable of proving that pain is not a modification of consciousness, for we can after all know when another is in pain.

It may, however, be claimed that meaning and understanding are in a special position, and that to treat them as resting even in part on inner states will give rise to a special sceptical problem not extending to all aspects of our mental life, just because meaning is essentially communicable. That is to presuppose that inner states of consciousness such as sensations, by contrast with meaning, are (or may be) incommunicable. Yet we can quite intelligibly describe our sensations to others. Indeed, as we have seen, the meaning even of a word like 'blue', a term primarily predicable of observable objects, is tightly linked to sensory appearance. If the content of consciousness is incommunicable, then the meaning of 'blue' is incommunicable. There could be no opening for scepticism about sensations or inner states in general which did not automatically extend to meaning. There can therefore be no specific reason for rejecting the natural presumption that our speaking with meaning (rather than parrot-fashion) and our understanding what another says (rather than just hearing it) are both intimately bound up with modifications of consciousness. That is not to agree with Locke that ideal understanding comprises a succession of images neatly corresponding to the subjects and predicates employed in the speech understood. The present point is simply that, since no one but a philosopher would deny or even doubt that the conscious state of someone who hears or reads a sentence with understanding is thereby significantly different from the conscious state of someone for whom the sentence has no meaning, any philosopher who does so had better have a more satisfactory argument than the one under consideration.

The second argument, which appeals to the premise that the criteria of meaning and understanding lie in performance, suffers from a similar inconsistency despite its allowing that there may be 'characteristic' accompaniments of understanding at the level of consciousness. For a precisely parallel argument could be advanced in relation to perception itself which would prove just as readily that states of consciousness are never anything but 'more or less characteristic accompaniments or manifestations' of perception, which perceivers could very well do without. It would admittedly be an odd sort of understanding speech which occurred without ever affecting or directing behaviour at all, but so too it would be an odd sort of perception. We should hardly believe someone's claim to see in the dark if they could neither give an account of what they allegedly saw nor avoid crashing into things with the rest of us. Yet that point as hardly proves that what we see, and how we see it, is ultimately a matter of verbal and other performance, or is independent of our conscious state at the time.

The discussion of perception and consciousness above, in [chapter 21](#) and elsewhere, contains some suggestions as to what is happening in these arguments about meaning and understanding. They do not involve an outright denial of consciousness, since the behaviourism involved is directed towards the explanation of the intentional content of the cognitive state in question, while consciousness is treated for the purpose of the argument as if it were a mere concomitant sequence of *qualia* without content. In this way behaviourism in

some form can coexist with a sensationist or imagist theory of consciousness. Such a combination, indeed, has constituted one of the more important philosophies of mind in this century, finding perhaps its fullest or most rigorous expression in Russell's *Analysis of Mind*, but often seeming to play a role in the subtler, less tangible, arguments of Wittgenstein, Gilbert Ryle, Elizabeth Anscombe and others.

Russell was persuaded by the behaviourist J.B. Watson that 'images are not logically necessary to account for observable behaviour', including 'intelligent speech'. Yet he could not accept that images do not occur, or that 'mental content' (by which Russell did not mean what has been called 'intentional content' in the present work, but rather consciousness as supposed to be constituted by images) does not commonly play some role in giving rise to behaviour. The role Russell ascribed to consciousness is explained by his theory of meaning. The core of the theory is that a symbol has meaning because of its associations, partly with other symbols but essentially with behaviour. The meaning of the word 'motor', Russell believed, must ultimately rest on such facts as that the word 'motor' can make you leap aside, just as the motor can.¹²⁹ Images enter into the theory not as meanings, but as symbols. An image of a shoe, like a picture of a shoe, is a 'natural' symbol for a shoe, but there is no necessity in the relation. The meaning of the image lies not in resemblance or correspondence, which may only be rough, but, exactly like the meaning of a word, in its associations: 'images may stimulate desire almost as strongly as the objects they represent'.¹³⁰ Thus, Russell asserted,

the question as to what the mental content 'means' (in case it means anything) is one which cannot be settled by merely explaining the intrinsic character of the mental content, but only by knowing its causal connexions in the case of the person concerned.¹³¹

The same image may have distinct meanings even for the same person, since different associations, with words and with behaviour, may be operative in different contexts. Or it may have no meaning at all. Only in a special context would a word or sentence have the same meaning as an image.

Arguments in some respects interestingly like Russell's are still in use, although those who advance them tend to see their origin in Wittgenstein.¹³² For example, Anscombe has applied the principle that modes of consciousness have no meaning in their own right to the task of proving that the intention of someone in acting has nothing to do with his conscious state at the time. According to one of her arguments,

If intention is an interior movement, it would appear that we can choose to have a certain intention and not another, just by, e.g., saying within ourselves: 'What I mean to be doing is earning my living, and not

poisoning the household...’ The idea that one can determine one’s intentions by making such a little speech to oneself is obviously bosh.¹³³

The reason is that

you cannot take any performance (even an interior performance) as itself an act of intention; for if you describe a performance, the fact that it has taken place is not a proof of intention; words for example may occur in somebody’s mind without his meaning them.¹³⁴

More recently Putnam has made a similar claim in his restatement of what he takes to be the Wittgensteinian argument about meaning. Meaning cannot consist in conscious thoughts because the question arises as to the meaning of any items in consciousness themselves. For example, someone might ‘mentally rehearse’ a long sequence of Japanese sentences without understanding them, a possibility which allegedly ‘shows the futility of a phenomenological approach to the problem of understanding’. Whatever ‘introspectible quality’ we suppose to accompany all actual cases of understanding, it is logically possible that the quality should be present in one who does not understand. On the other hand, Putnam argues, there might be someone who ‘speaks perfectly good English’ without any ‘interior monologue’ at all: ‘he never thinks (in words, images, etc.) when he is not speaking out loud; nor does anything “go through his mind” when he speaks out loud, except that (of course) he hears his own voice speaking, and has the usual sense-impressions from his surroundings, plus a general “feeling of understanding”’. As Putnam remarks with some complacency, the conclusion that what goes on in consciousness is irrelevant to understanding is ‘one of the most remarkable conclusions in the history of philosophy’.¹³⁵ Yet we are left to wonder why his English-speaker is allowed ‘of course’ to have sensations and ‘feelings’. If we suppose that he lacks them, then still, *ex hypothesi*, nothing has changed to the onlooker, who must of course rely on performance—on what else? If it is said that such a supposition replaces the person of the story with a mere robot, then it should be seen that the argument for treating Putnam’s semi-robot (or robot with sensations) as an intelligent speaker of English is no stronger than the argument for treating a totally insensate robot as an intelligent speaker of English.

To reject the ‘Wittgensteinian’ argument is not, of course, to revive the broadly Lockean view, against which that argument is directed: that when a number of people understand a word or sentence in the same sense, then precisely the same thing must occur in the consciousness of all of them, such as a particular abstract idea or string of ideas. The logical criticisms considered in chapters 28 and 30, above, are enough to make that view untenable, and more are available. What is here objected to is, first, the sweeping dismissal of consciousness from the account of linguistic understanding (as from the accounts of understanding in general, and of other mental states, such as intention) and,

second, the assumption by which that conclusion is achieved, that consciousness consists in nothing but a string of sensations, images, verbal images and 'introspectible qualities', together with 'feelings' of desire for something, of certainty, of understanding and so forth, the latter conceived of as quasi-sensations. The net effect of this assumption, oddly enough, is not easily distinguishable from an imagist theory of consciousness. Yet it seems to spring not so much from the imagist tradition (except in Russell's case) as from the assumption that it must be possible to describe any aspect of 'what is going on in consciousness' in neutral terms, as it is intrinsically and in itself, without reference to its intentional content. 'Conscious desire' thereby becomes a blank 'feeling' conceived of as a quasi-physical sensation, neither necessary nor sufficient for real desire for something. That sensations and images themselves must have intentional content is either implicitly denied or conveniently overlooked. Any 'conscious thought', whatever its type, is reduced to a pictorial or verbal image. But an image of a burning house, although it has content, cannot as such have the content of the thought that my house is burning down, or of the intention to start a fire. Similarly a verbal image, although it has an object, namely the physical sentence as uttered or written, inevitably fails to contain in itself the intentional content of the thought expressed, just because any sentence might be imagined without being understood.

It is quite unsurprising if such an argument drives what is important to thought beyond the scope of consciousness, since what is really at issue, the boundary of the content of consciousness, is merely presupposed. Criticism has already been offered above of a range of arguments which would do just the same for sensations themselves. If it is assumed that the intrinsic character of sensations can be fully described without reference to their intentional content, it will obviously follow that their content is something added by an extrinsic relation, whether to their causes or to ensuing belief and, ultimately, to behaviour. If that assumption were correct, it would indeed 'show the futility of the phenomenological approach' to the problem of the intentional content of sensation. Yet as I have argued at length and from a variety of directions, for all sorts of theoretical and common-sense reasons sensations must be regarded as intrinsically intentional, as being in themselves presentations of spatial reality in consciousness. Such an approach, for example, allows an intelligible explanation of the perspicuousness of perceptual knowledge, and of the intuitively attractive, philosophically explanatory distinction between primary and secondary knowledge. The present suggestion is that linguistic understanding too is characteristically, in primary or paradigm cases, perspicuous, and that it is not surprising that its perspicuity, like the perspicuity of *a priori* knowledge, cannot be explained in terms of sensations, sensory images, feelings and the like.

It will be seen that the Wittgensteinian argument merely begs the main question. Can the intrinsic intentionality of consciousness be supposed to go *beyond* what falls into the categories of 'sensation', 'feeling' or 'image'? There is no obvious reason to suppose that it cannot. On the contrary, as Putnam

proudly records, we feel intuitive astonishment at the doctrine which tells us that understanding the sentences we are reading involves, as such, no effect on our conscious state at the time. That, presumably, is why the argument for that doctrine is so dependent on an instilled sense of the depth of the misconceptions into which we have fallen, and from which we need to be rescued.

Nevertheless it can seem strangely difficult to say what kind of effect in consciousness it is which is involved in understanding (or at least in primary understanding). In order to resolve some of the difficulty it may be best to begin with analogies at some distance from the topic of language. Take, for example, a simple case of memory. A man who embraces the long-lost sister whom he has met by chance does so *because* he consciously recognizes her. It is entirely unnatural to offer to analyse what happens solely in terms of his seeing her and consequently acquiring such attributes as the disposition to shout her name, whether out loud or to himself. We normally employ the notion of recognition precisely to do justice to the exercise of a cognitive capacity at the level of consciousness. Yet the Wittgensteinian strategy is here to argue that, if there were a conscious element in this case of recognition, over and above bare sensation, which constituted a necessary condition of recognition, then that element would have to be independently describable as some further image or feeling, e.g. an image of the sister when young, or a feeling of familiarity or of nostalgia. It is then easily shown that such an image or feeling is neither a necessary nor sufficient condition of recognition. Yet, as the comparison with perception proves, there is in fact no good reason why the element in consciousness in the man's recognition of his sister should be describable *independently* of the concept of recognition, memory or other cognate concepts. For in the case of perception the necessary conscious element in the perception of a blue circle, i.e. the visual sensation, is not after all describable independently of the notion of the perception of a blue circle. To describe it as a sensation or image or idea or impression or sense-datum of a blue circle is a palpable failure to perform that task. Indeed one might locate what is primarily wrong with the traditional 'sense-datum' approach in its assumption that the task of independent description *has* been satisfactorily performed: as if, having independently identified blue circular ideas or sense-data or *qualia* or 'raw feels', we could *then* raise the question whether an awareness of these things might not have something to do with the perception of physical blue circles.

To return to our example of recognition, it need hardly be said that what happens in the man's consciousness cannot constitute a *sufficient* condition for successful recognition, since he might be mistaken. Recognition is not the exercise of an infallible cognitive faculty, and there is an objective, external condition of recognition as of seeing and knowing generally. Nor is it possible that a man should recognize Mary, or recognize someone as Mary, whatever goes on in his consciousness, if he lacks all previous knowledge of Mary. The present recognitional state must be due to past experience of Mary (or of a photograph of Mary, or the like) through the operation of the mechanism of

memory. Yet such objective conditions are not sufficient either, for, like primary memory generally, primary recognition is 'perspicuous'. If a man recognizes someone as his sister, there is no mystery to him as to why he takes her to be his sister, even if he cannot explain in words what has made his sister memorable. We can imagine another possibility: the man does not recognize a certain woman as his sister, but nevertheless has a firm belief or hunch that that is what she is. If the belief is true and not accidentally so (i.e. it is due to appropriate past experience through a dependable memory-mechanism), then we might talk of unconscious recognition and memory. The case is like those cases of 'perception without sensation' discussed in chapters 15 and 19. It is necessarily an atypical or secondary case of recognition, as they are of perception. Just as the supposition that all our perceptual knowledge should be secondary is absurd, so is the supposition that all recognition should be secondary.

All these considerations indicate that we cannot explain recognition, any more than we can explain sense-perception or, indeed, knowledge, unless we grant that it typically involves consciousness. At the same time the Wittgensteinian argument, if we stand it on its head, shows that the element in consciousness which is a necessary condition of primary recognition cannot be construed neutrally, as for example Russell tried to construe it, in terms of images and quasi-sensory 'feelings of familiarity' or pastness. In other words, we have identified an element in consciousness which, like, sensation, is irreducibly or intrinsically intentional, but whose intentional content goes beyond the content of any sensations or images. That is not, however, to suppose that that element could exist in a pure state, in the absence of any sensation or image, like a Cartesian idea of pure intellect. Recognition necessarily involves sensation and not simply as a cause, for it is the perceived object itself which is imbued with significance or import, that significance which it will lack for one who sees it just as well and clearly, but who fails to recognize or place it. Nor should we be surprised that rather different things happen at the level of consciousness in different cases or kinds of recognition, a point of a sort frequently made in the Wittgensteinian dialectic. A familiar face recognized in a strange crowd will thereby catch particular attention, but a familiar object recognized in its usual place in the house will thereby escape it. The resemblance between the two cases lies in what is captured by the notion of recognition itself, not in the vivacity of the experience or in the element of attention or, for that matter, in its effect on 'performance' (so much the worse for behaviourism). There is likely, it is true, to be a certain complex but systematic relationship between the content of consciousness and behaviour, for after all the former tends to determine the latter even in the severely insane. But the relationship does not itself constitute what is special about recognition.

Consider another sort of case. Suppose that I see a type of artifact for the first time and realize (or think that I realize) what it is for. What happens in consciousness? A Wittgensteinian might well have ready the familiar list of possibilities. Perhaps a feeling of puzzlement gives place to a feeling of relief.

Perhaps I imagine using the object to put on a shoe. Perhaps the sentence 'It's for putting on shoes' passes before my mind. Or perhaps I simply utter that sentence (and, of course, hear myself uttering it). Or perhaps *nothing* happens in consciousness beyond my sensations of the object, but when the time comes I simply use the object as a shoehorn. And so forth. What happens in consciousness, it is suggested, is simply a colourful accompaniment: what matters is performance. The general argument fits straight off its peg in the *Philosophical Investigations*. In reply we may say that *realization* does have objective conditions: it must be of what is true and must constitute the operation of appropriate cognitive mechanisms. Yet to make it something other than an event in consciousness would not only be a blatant paradox, but would make realization entirely mysterious. It is quite true that it is not enough, in our example, to form an image of someone's using the artifact to put on a shoe. For I might whimsically imagine easing on a shoe with the salad spoon without in the least supposing that the spoon is for putting on shoes. Yet the fact is that, in consciously realizing what the artifact is for, we do somehow attach the content 'for putting on shoes' to this perceived object. Moreover, while an image of the act may be both inessential and insufficient, it is surely no accident if, in realizing an object's function, I do have some kind of partial, scrappy, motor image of my using it for that function, or a visual image of somebody else doing so. Such an image may serve as the medium through which the content of the realization, which here necessarily goes beyond what is sensory or imageable, is partially expressed: it serves, so to speak, as a hook on which further import is hung. It is inessential in so far as the object as perceived may itself be sufficient for the purpose.

That kind of case offers a promising model for linguistic meaning and understanding, if we suppose that heard, seen, spoken and imagined sentences are other such hooks, constructible only after intensive training (with or without an innate head-start), but essential for certain sorts of sophisticated intentional content. Understanding a sentence certainly seems like a case of recognition, and it is also like the understanding by which, for example, we can just 'see' how a mechanism, actually or in plan before us, will work.

These proposals raise the very large problem of what a mental 'image' is anyway. The issue here may be confused by the quite general use of the English word 'imagine' to stand in for 'suppose' or 'conceive', a use for which philosophical imagism must itself be historically responsible. 'Images' and 'imagery', however, have a natural and narrower use for the phenomena of what has been called above 'sensory imagination'. The present question concerns those phenomena.

In a certain class of cases a sensory image seems simply to be an element or aspect of our sensory experience itself which is employed or taken in a certain way. In other words, it is a sensation bearing a certain sort of import or content which goes beyond the intentional content of the sensation as such. In these cases the objects of imagination are, in another guise, actual objects of

perception: for example, I may employ what I see as I gaze into the fire as a means of imagining a face. In so far as I am contemplating an image, it is located where it appears, in the fire itself. I am not of course misperceiving the fire as a face, nor is what I am doing just like seeing a photograph as a picture of a face. The latter is a commonplace achievement of vision (however founded on earlier experience or training), whereas in imagination I *fill out* or *complete* what I literally see, as a child may imbue a doll with life, or turn a run across the playground into a cavalry charge.

At the other extreme, there seem to be images less directly associated with actual perception. Memory images may be like this, as pathological cases of vivid recall illustrate in an exaggerated way, involving what has been called a 'doubling of consciousness'.¹³⁶ Even in an ordinary case of recall it would seem possible, for example, to run through a tune in one's head without the aid of existing or induced sensory stimulus, however natural and helpful it may in fact be to breathe in appropriate rhythm, move the larynx, nod the head and so forth in doing so. It seems equally possible to form a visual image of the breakfast I had this morning without employing any of my current visual sensations. It is no doubt such cases which have supplied philosophers with their paradigmatic operations of the imagination, conceived of as a faculty separate from the senses, with its own organ or seat. Yet the characteristic entanglement of sensation and imagery would seem worth bearing always in mind. Notoriously an actual sensible event such as the ringing of a bell can become incorporated in a dream, and the same may go for the dreamer's sensations of his actual bodily movements. So too, in waking life, someone imagining himself taking part in a fight may 'hear' cries, including his own, and 'feel' the blows he makes or receives, with the aid of scrappy auditory or motor sensations.

The imagination of pain typically takes a similar form. It is possible to pick on some bodily sensation and, so to speak, transform it into pain in the imagination. It may help to shift slightly, expanding our motor sensations into images of appropriate writhings. The emotional context—horror, sympathy, fear—may contribute to the 'vivacity' of the image, or indeed may be a component of it, while the image itself may add to our distress. As Berkeley and other imagists have remarked, real pain is very different from the idea of pain. Yet the imagination of pain does characteristically involve bodily sensation. Or take another kind of case. I look at the pen in my hand and vividly imagine it turning over. I can 'just see' or 'just feel' it turning, although it does not perceptibly turn. It would be an unrealistic abstraction wholly to separate my 'image' of the pen from my sensation of the pen, or the object of imagination from the object of sensation, as if they occupied different mental spaces. It is just that pen, as I see and feel it, upon which my imagination is focused and which I imagine to be turning over. Perhaps in doing so I move it very slightly, its movement serving in much the same way as a very slight pain can promote or, rather, help to constitute the image of a severe one.

In other cases the elements of sensation employed in imagination seem to consist less in the ordinary sensation of things around us, or of our own bodies, than in sensory 'noise' which is disregarded and shut out in normal perceptual experience, but which can become obtrusive in special circumstances, as in periods of quiet or repose. Such 'noise' is not itself without intentional content, but its content is not straight-forwardly integrated with normal perceptual content. For example, as I move my gaze across the window I am aware of dark, transparent patches or spots moving on my visual field which do not appear to be 'out there', but through which I see the window and what lies beyond it. So too I can become aware of a continuous rustling which appears rather to be in my head than in the room. Such phenomena may perhaps lead to illusion, the appearance of a cry or of a flying saucer. By the same token it constitutes material on which the imagination can build. When I shut my eyes, the fluctuating patches or patterns which emerge on a visual field of indeterminate depth can, it seems, be marshalled as scrappy representations of something more substantial. The imagined object does not appear outside us, but it is at the same time 'before my eyes' and 'in my head'. Something similar is true of after-images. To have an after-image is not to imagine something, but both after-images and some of the images employed in imagination are by-products of the mechanism of vision with this kind of ambiguous content.

It is not perhaps particularly surprising that some psychologists who take imagery seriously enough to investigate it have proposed that the psychological process of imagination is a sort of rehearsal of the process of sensation, involving some part of the mechanism of perception itself rather than an independent mechanism of its own. Such a conclusion is consonant not only with the datum that the mode of representation of images is characteristically scrappy and suggestive as compared with the mode of representation of sensations, but also with the close link between images and sensations from the subjective point of view. Certainly, in so far as sensory images are constituted by sensations or even by sensory 'noise', they are in some way integrated within the unitary sense-field; even if, like after-images or spots before the eyes, they are sometimes imperfectly or ambiguously integrated with the more objective or straightforward deliverances of the senses. Even those more ethereal visual and auditory images which seem less closely involved with actual sensation are naturally described as 'in the head', no doubt because that is where we see and hear things from: the presentation of the objects of the imagination, despite differences, still has that much in common with the presentation of the objects of sense.

These are questions calling for systematic empirical investigation, but that is not to say that it might be discovered that images do not occur at all. Those who see their very existence as problematic (and it may be that images currently have more enemies than friends among academic philosophers) are for the most part motivated by the general prejudice against consciousness or subjectivity. The present argument is concerned only to justify finding room in an account of our experience for something at least a bit like the traditional conception of

imagination as 'decaying sense'. There may be myths about images, but images are not themselves a myth.

Yet, although images surely do deserve a place in our account of thought, the doctrine which has here been called 'imagism' is surely false. Something approaching it seems to be true, in so far as it is plausible that all conscious thought involves sensations or images of some kind. It does not seem possible to make sense of the notion of a purely intellectual process of conscious thought totally disconnected from everything sensory, including verbal or otherwise symbolic images. No thinking thing could exist which was incapable of sense-perception. Yet arguments like Russell's and Wittgenstein's, despite themselves, undermine the stronger form of imagism according to which sensations and images (including image-like 'feelings' and the rest) exhaust the contents of consciousness. For present purposes Locke's imagism can be treated as a version of this stronger doctrine. For him, it is true, consciousness comprises, as well as sensory ideas, such 'operations' as distinguishing, attending, comparing and abstracting, not to speak of our taking ideas as signs and conjoining them in propositions and chains of reasoning. Yet sensory ideas in their role as signs of their causes are, on his account, the bearers of all the intentional content of thought, with the exception of the causality of which we are supposed to be immediately aware in sense-perception. In the present chapter, in contrast, it has been argued that, even if all conscious thought involves sensations or images, the intrinsic intentional content of conscious thought typically goes beyond the intrinsic intentional content of the sensations and images involved in it. Our conscious realization that the person we are talking to is hostile to us essentially involves sensations, but the content of our consciousness goes beyond the content of the sensations in some way which cannot be explained by postulating (perhaps in the face of experience) concomitant *images* of one sort or another. It seems to be in a similar way that the conscious state of one who hears and understands a sentence will thereby be richer in content than the conscious state of one who hears the sentence without understanding, even if the latter happens to be having images appropriate to the subject-matter of the sentence. Like all cognition, linguistic understanding occurs only when the construal placed on the sentence is the right one, and when its rightness is non-accidental, resulting from the operation of appropriate capacities. But the construal will also characteristically be 'perspicuous': it will not arise, from the subjective point of view, as a mere hunch. It is this last condition of primary understanding which the Wittgensteinian argument, like so much modern philosophy, simply rejects; and it is because of the necessity that this condition should be fulfilled that primary understanding is, after all, essentially a conscious process.

Necessity, reason and language

The explanation of *a priori* knowledge which is possibly still the most popular among analytic philosophers is that it is at bottom linguistic, in some sense knowledge of linguistic rules and nothing more. It would seem to follow directly that the ability to reason deductively and to spot any sort of logical or conceptual relationship is to be explained as a form of linguistic competence, and that rational 'intuition', in Locke's sense, can be nothing but a type of linguistic intuition. Yet there are notorious difficulties for such a view, starting with Descartes' response to Hobbes, that 'a Frenchman and a German are able to reason in exactly the same way about the same things'.¹³⁷ There seems to be a clear distinction between understanding an argument as a bit of (say) English, and understanding or getting the point of the reasoning. Indeed a failure to reason correctly seems so far from being a sign that the sentences in question are not fully understood, that, unless they are fully understood, the case cannot be one of invalid reasoning. The linguistic competence required for an understanding of the sentences seems to be one thing, the reason or intelligence required to follow or assess the argument quite another. The ability to construct proofs is something else which is naturally included under a facility with ideas rather than with language.

There are, of course, responses to these objections to the linguistic theory of *a priori* truth. It may be said that, while the ability to think up the steps of a proof count as a part of rationality or intelligence and cannot simply be brought under the aegis of linguistic competence, yet the only factors determining the validity or 'rationality' of any step, once made, are meaning-rules. Hence either to make or to accept the step non-accidentally is in itself simply to demonstrate possession of these rules. Again, it might be admitted that truths of reason need not be self-evident or obvious if the linguistic facts they illustrate are complex; but it can be argued that it is only for that reason that understanding premise and conclusion is not sufficient for seeing the logical connection between them. The kind of insight and intuition possessed by one who does see the connection (it may be argued) is still insight into the formal structure of language, and support for it must come from the making explicit a series of steps each of which cannot be denied by anyone who understands the language. Producing these steps may

require considerable intelligence, but following them does not for anyone who knows their meaning. Moreover a Frenchman and a German may have the same insight, because the insight is into linguistic rules which different languages may have in common.

Nevertheless, except for the dogmatic reductionist, the problem remains. 'All brothers are male' may seem nothing but the expression or manifestation of an arbitrary linguistic rule, but the intelligible laws of logic appear in relation to language as prior constraints rather than as mere constitutive rules. On the other hand, to treat them as entirely independent of language seems absurd, for no one could syllogize without sentences. We need a route out of this *impasse*, and a promising policy is to reflect on the role of rationality or, less portentously, of general intelligence in the acquisition and employment of language. For in an explication of *a priori* insight reference to semantic rules may well seem simply circular. Language is not, after all, a pointless noise-game or conventional ritual to be learnt by rote or mere habituation. The distinction between following a proof and understanding the language in which it is expressed may not be nearly so sharp as the traditional critic of the linguistic theory suggests. That is not because there is no more to following the proof than understanding the English, for that way of putting it implies that 'reason' is here a superfluous faculty, or has been explained away as linguistic competence. It is rather that intelligent comprehension of a more or less difficult issue or technique may necessarily be involved in understanding a piece of language, a point which leaves reason enthroned but which does not imply that the issue could equally well have been comprehended without the language.

It is obvious that the mastery of technical terminology commonly requires an intelligent grasp of abstract subject-matter. No one could understand what lawyers currently mean by 'negligence', or philosophers once meant by 'knowledge' or 'substance', who is intellectually incapable of following the arguments which have historically determined that meaning. Yet there is neither a sharp nor clearly significant distinction between the technical and the everyday, the abstruse and the ordinary. To instruct a child in familiar points of language may sometimes seem to be no more than to teach the conventional etiquette of the language-game. Yet the explanation even of such a thing as the difference between the perfect and pluperfect tenses, however informally it is done, involves not merely an exposition or illustration of syntactic rules but an invitation to grasp the point or significance of the difference between using the one tense and using the other. Another example lies at the heart of mathematics. Intelligence is required to grasp even the first step in arithmetic, the technique of counting, but that technique would be impossible without language. Consequently understanding a language is closer than may be supposed to the understanding which involves 'seeing why', and which is connected with justification and with a more profound kind of knowledge than mere knowledge how to play a rule-governed game with sounds or marks.

It may here be helpful to consider the modern theory of innate knowledge advanced by Noam Chomsky and his followers, a theory which specifically concerns the role of general intelligence in language learning. Chomsky was originally reacting above all to orthodox behaviourist learning theory, according to which language acquisition is in principle explicable entirely in terms of conditioned responses, i.e. crudely speaking, in terms of a general capacity to be habituated. For the behaviouristic psychologist such a capacity constitutes, if anything does, 'intelligence', and is present in human beings to a high degree. Chomsky argued that this theoretical resource is quite insufficient to explain language learning, basing his argument on the sheer complexity of the rules uncovered by transformational linguistics. It is, he claimed, necessary to postulate an innate foundation for our propensity to pick up language as rapidly as we do. At the same time the model which emerges from Chomsky's argument assigns a certain role to intelligence. The proposal seems to be that the child unconsciously approaches linguistic phenomena rather like a scientist, advancing by speculative hypotheses which are suggested by the data and open to confirmation, rejection or modification in the light of experience. Chomsky argues, however, that, given the complexity of language, unless we also suppose a set of innate 'organizing principles' or 'restrictive conditions' on the form that such unconscious hypotheses could take, it remains incredible that the data should provide the basis for a speculative hypothesis with more than the slightest chance of being right. These innate constraints are constituted by the common 'deep structures' discernible in all natural languages by the transformational grammarian.

Now, as some have pointed out in criticism of Chomsky, the postulation of innate structure does not in itself constitute a very radical departure from behaviourist doctrine. A dispositionalist or stimulus-response account of linguistic capacity would not be seriously upset by the news that some high-level linguistic dispositions are, doubtless for good biological reasons, inborn. Nevertheless Chomsky's conception of deep structure as a constraint on a general speculative intelligence does seem opposed to behaviourist psychology if only because of his conception of speculative intelligence. But another criticism is that, taking at least his earlier arguments overall, there seems to be a remarkable inconsistency in his conception of intelligence. For he argued not only that the capacity to learn language is relatively independent of intelligence, but that the study of 'deep structure' or 'universal grammar' is concerned with 'the nature of human intellectual capacities'. According to this second line of thought, so far is the innate possession of universal grammar from being independent of intelligence, that the linguist studying universal grammar is specifically 'trying to establish certain general properties of human intelligence'.¹³⁸ In fact certain assumptions which Chomsky makes about what he calls the 'semantic component' of grammar add up to the belief that one object of psycholinguistics is the form of the innate mechanisms which determine the way we reason in language, and what we find reasonable. That is why his follower, J.J. Katz, could argue that the

systematic and empirical study of languages should enable us to resolve conceptual disputes,¹³⁹ in particular philosophical disputes. Chomsky himself has made the claim that his theory incorporates the classical theory of innate ideas, the chief point of which was to explain our knowledge of truths of reason, i.e. to explain *a priori* rationality. Hence 'intelligence' is on the one hand directly opposed to innate semantic principles, but on the other is taken to be embodied in those principles. As we shall see, if this curious duality is eliminated and the two conceptions of intelligence are drawn together, the thesis of innateness becomes superfluous.

Chomsky's empirical argument has, for the most part, avoided the 'semantic' component of grammar in favour of the other components, the 'phonological' and 'syntactic'. It starts from such deviant sentences as 'What boy did he believe the claim that John made about?', an example which Chomsky locates in the class of those that are 'quite impossible, although it would be clear what they meant, were they grammatically permissible'. Here most would have an immediate 'intuition' in something close to the modern conversational, rather than the traditional philosophical sense: as it were, a dependable sense of wrongness, without understanding why. How we know the sentence is wrong is at first obscure, a thing to be made plain by a grammatical explanation. If the explanation is complex, involving underlying principles common to all languages, then there appears to be at least an opening for the hypothesis of innate syntactic rules. Yet such a hypothesis has nothing to do with the structure of human intelligence or rationality. That is because rules for pronominalization, the reference of relative clauses and so forth, while they might seem in some sense natural, can hardly be rational or 'evident'. We cannot see why one such rule must hold rather than another, or even why there should not be inconsistencies in idiom. Similarly, if there were indubitably a common phonological structure to all languages which was demonstrably innate, the discovery could hardly be characterized as one about the human 'intellect' in any ordinary sense. That would be so even if knowledge of the structure helped to explain, for example, certain difficulties in learning language among certain brain-damaged but intelligent children. We might compare such difficulties with dyslexia, which is a recognitional, not an intellectual limitation.

The association of innate principles with human reason must therefore rest on the possibility of extending Chomsky's argument to the so-called 'semantic component' of grammar. But it should already be clear that such an extension may well be impossible in principle. For a semantically deviant sentence, e.g. 'If John is taller than Mary, then Mary is taller than John', cannot supply the same material for Chomsky's argument as the merely syntactically deviant sentence just because our sense of wrongness is quite different in the two cases. If we talk of 'intuition' in the semantic context the word is used more or less as Locke used it, for recognition of the self-evident. Since it is evident that what the sentence would assert is impossible (in other words, we immediately 'see why' the sentence cannot be true) there is not the same foothold in the content of our

knowledge, in respect of that feature of the sentence which makes it unacceptable, for the argument that we are being implicitly guided by a possibly complex system of rules. This 'evidence' peculiar to semantic contexts should not be confused with subjective conviction, for our reaction to the syntactically deviant sentence may be just as unqualified and convinced. The judgement of merely syntactic deviance is the exercise of a skill or competence, an ear for the linguistically correct, whereas the judgement of semantic deviance is an exercise of intelligence involving understanding. It can therefore be objected to the argument for innate semantic principles that it requires us to treat the 'intuitions' of reason as on the same level as linguistic or, as it might be better to say, merely linguistic intuitions. It supposes that what is perspicuous requires the same sort of explanation as what is not perspicuous. That objection, of course, is closely related to Locke's objection to the classical theory of innate knowledge.

A fairly strong distinction between grammar and logic, between the conceptual and the merely grammatical, is thus not without foundation. The judgement that a sentence is grammatical can be distinguished from the judgement that it 'makes sense'. The deviant sentence 'John is more good at puzzles than Bill is good at puzzles', while execrable English, 'makes sense' as the Escher-like 'John is better at puzzles than Bill is better at puzzles' does not. The argument for innate semantic structure presupposes that behind the judgement that a sentence does or does not 'make sense' lie implicitly accepted linguistic rules and nothing else. Yet behind the judgement that the first, ungrammatical sentence 'makes sense' there lies recognition of the simple possibility of comparing levels of excellence, a possibility no more due to grammar than the possibility of any quantitative comparison. If it seems empty or mysterious to talk of such recognition, consider 'John is better at reading by more than Bill is better at writing', 'John is better at reading by more than Bill is better at reading' and 'John is more better at reading than Bill'. Someone hearing these sentences might easily, at first response, think that all are deviant: but only the last of the three is in fact deviant, and all can be regarded as 'making sense'. To grasp that the first is not deviant it is necessary to grasp the possibility of comparing the degree to which John excels at one accomplishment with the degree to which Bill excels at another. With the second sentence, we must grasp the possibility of comparing the degree to which John's ability at reading excels another of his abilities with the corresponding degree in the case of Bill. Since it is easy to suppose that someone should, through lack or temporary failure of intelligence, be incapable of grasping or slow to grasp these possibilities, it is more than plausible that in interpreting what is said to them children have in their armoury not only linguistic intuitions, however inborn or acquired, but also their general intelligence or capacity for seeing what 'makes sense'.

The specific context of utterance may be important to the exercise of this capacity, as a rather different kind of case will illustrate. Katz argues that the ordinary English imperative, for example, 'Help the man', must be grammatically conceived of as being derived from an underlying structure or

'phrase-marker' with 'you' as the subject. This 'unobservable' feature contributes to the meaning, so his argument continues, and so must somehow be known by one who knows that meaning. Since the feature is unobservable, it cannot be allowed that a conditioning process leads the learner so readily to predict, for example, that only 'will you?' or 'yourself' will ever be correct after an imperative, never 'himself', 'will he?' and so on. Consequently knowledge of the feature must be innate.¹⁴⁰

What, if anything, is wrong with this surprising argument? The principle that 'you' is the subject of an imperative has a decidedly *a priori* air. A child's implicit realization that the imperative 'jump' has 'you' as unheard grammatical subject is not distinct from his realization (given prior knowledge of the sense of the verb *jump*) that, when authority says 'Jump!', the person addressed is expected to jump. But the latter realization is likely to come when he sees the others jump, the impatience of the speaker directed towards non-jumpers and so on, provided that he has the intelligence to propose such a mildly creative interpretation of the whole situation. We can then understand why the sentence 'Dress yourself' should appear unexceptionable, whereas, if he has a grasp of reflexives from other contexts, 'Dress herself' should appear either not to make sense or not to be an imperative (or to introduce a novel use of 'herself'). Thus there is no need or even reason to postulate, as Katz does, that the child has unconsciously derived the final phrase-marker, by implicitly recognized principles of derivation, from an underlying phrase-marker constructed on implicitly known principles of universal grammar. Such a portentous interpretation seems, indeed, to rely on a quite unrealistic abstraction from the situations in which significance is grasped. The 'input' upon which the child's intelligence can work is too narrowly conceived, and only for that reason can it seem that the child must be directed by innate semantic-syntactic principles.

Chomsky's neglect of context is embodied in his model for the role of speculative intelligence in language acquisition, for that model would better fit the method by which an observer ignorant of chess might attempt the abstract task of systematizing the moves of chess-players in order to arrive at the fixed and arbitrary rules by which they play. But chess is not like language precisely in that the determinate rules of chess have nothing directly to do with the point of playing. Chomsky seems to assume that the child's intelligence is immediately, if unconsciously directed towards uncovering a pattern in the narrowly defined input, the presented sequence of sounds. On a better model, intelligence is also continually, and more consciously, directed towards something other than purely linguistic phenomena and their internal structure: namely, towards the point, in the total context, of this or that piece of language or 'language-game'. What this means can perhaps be made less obscure by considering the role of intelligence in abnormal uses or extensions of language. A metaphor or paradox achieves its force, to put it as neutrally as possible, by its employment of words out of their usual contexts, or in some sense by 'bending' standard usage. It seems to follow that habituation could not be enough, but intelligence is required, for the hearer

to get the point. Yet there is no significant difference in principle between coming to grasp an established use of language and getting the point of a novel one. For either, for example, we may need to grasp an explicit or implicit justification.

It should be helpful to return to a certain criticism of Locke's conception of language which was mentioned in [chapter 28](#), above. There it was suggested that in the case of vague predicates a perceived or remembered or imagined instance, however abstractly or partially considered, could not be supposed to contain within it some principle which determines the extent of the general predicate. All red things are alike, but differ even in their redness. They do not share a common feature, identical in each and every one of them, which determines that it is properly called 'red'. The role of 'use or consent' in classification cannot therefore be limited, as Locke limited it, to our linking a word with an already determinate feature or point of resemblance defining a class. Established convention or *ad hoc* stipulation must itself determine the boundary of the concept, and we normally rely simply on our judgements as speakers of the language when we decide what does and what does not satisfy a particular predicate. If we talk of perceiving in several things the common feature *redness*, we should realize that there would be no such determinate and identifiable 'feature' or point of resemblance unless some word had the meaning which 'red' has. Hence we cannot appeal to the evident presence or absence of such an antecedently identified feature in a dispute over a borderline case: we have rather to appeal to a notion of what would *properly* be called 'red'.

Despite lexicons, however, the rules for proper usage are not laid up in libraries, and our appeal is ultimately an appeal to further judgements as a user of the language. In other words, a disputant will appeal to supposed analogies between the case in question and other particular judgements which he assumes will be uncontroversial or generally agreed. Sometimes particular judgements may actually have been antecedently agreed between the disputants and so can serve as precedents. As John Wisdom in particular has pointed out,¹⁴¹ such an argument can be very like an argument in British common law, in that the appeal is often to particular cases rather than to principles, and, if to stated principles (as might occur with predicates more complex than 'red'), then to principles which rest on particular judgements and which themselves need interpretation in the light of particular judgements. There is no set of definitions or incontrovertible rules to which we can turn infallibly. One peculiarity of an argument in law, however, is that actual, precedent, formal judgements can have a special authority. To learn the extent of a predicate in common law, such as 'negligent' or 'vehicle', we must learn that such and such precedent cases were decided in such and such ways and (for what it is worth) for such and such given reasons. Common lawyers all work within the same range of cases, and at any one moment they are roughly analogous to a group of people who have agreed to settle disputes over the application of colour-words by reference to a set of

samples—although without prejudice to the possibility of challenging and rationalizing the samples.

Thus an appeal to proper usage is not an appeal to common usage, i.e. to any and every recorded instance of what people say and write, the first crude, unsifted material of lexicography. For there is something normative about such an appeal which presupposes that there is not just public usage, but a public standard. Its basis is not indiscriminating sociological observation, since it partly comprises, or is continuous with, an appeal to reason and consistency in judgement. The point is not readily illustrated by colour-predicates, if only because there is not as much scope for rational argument over calling something 'red' as there is over calling something 'negligent'. Yet with respect to other ordinary vague predicates, 'game', 'tool', 'democracy', 'deception' and so forth, there may even be greater room for disagreement than in the legal case. There may not even be a rough and elastic set of actual paradigms, as there is in common law, consciously and deliberately shared by all who know the meaning of the word. It is therefore characteristic, and not an aberration, that those who well know the meaning of these predicates should nevertheless sometimes disagree, and disagree rationally, over their application, even when no relevant circumstance is unknown or in dispute. Rather than saying that common usage supplies an authoritative and determinate boundary for the extension of a word, we might say that that usage supplies both the possibility of drawing a boundary and more or less firm indications of where the boundary might properly be drawn, but that this material for judgement is in principle open to different intelligent and argued interpretations. To endeavour to strip language of this openness and, in some sense, elasticity, through the relentless pursuit of precision and determinacy, is to attempt (as is no doubt impossible) to remove from it what is essential to its role as the tool of reason. It is a part of the nature of language that even established ways of talking can be rationally criticized, and that language can be rationally modified or extended from within, whether permanently, as perhaps Freud has extended it, or *ad hoc* to make a point.

Here we might remember that philosophical disputes themselves comprise a significant class of *a priori* disputes, a class which, according to Katz, transformational grammar promises to resolve. Philosophy, we are told, has been 'puzzling about words for over two thousand years', and we shall remain condemned to 'endless quibbling' in the absence of a theory of natural language that will enable us to 'extrapolate from the clear cases in such a fashion as to extend its generalizations to the unclear cases, thereby utilizing our strong intuitions about clear cases to compensate for our weak intuitions about unclear ones'.¹⁴² For Katz all relevant 'intuitions' are linguistic, and in this he stays close to the presuppositions of traditional analytic philosophy. Yet the reason why philosophical issues are liable to seem intransigent is not that philosophers have underestimated the importance of generalization (of which, indeed, they might be thought inordinately fond), but because they must, in the nature of the case, appeal to a conception of what 'makes sense' rather than to blank principles of

grammar. And even when wrong, what they say can have a point, or satisfy some intelligent judgements, if not all. The jaundiced view of the history of philosophy as a history of puzzling and quibbling over words is misconceived, drawing its rationalization from blind alleys or swings of the pendulum in the past (not to speak of bad history in the present), but rebutted by the role of philosophy in our increasing grasp of the world in which we live and of the problems which face us. The enormous intellectual upheaval which constituted the evolution of modern conceptions of nature and natural science (and of which the *Essay* is an unsurpassed manifestation) is itself a shining counter-example to the linguistic theory of the *a priori*. With due respect to Chomsky and his followers, it is not clear what contribution transformational grammar might have made to that tremendous advance in *a priori* understanding. But it is not particularly surprising that the century of linguistic philosophy should be a time of systematic denigration of previous philosophical achievement.

The suggestion that even in our intuitions of evident necessity we are characteristically being guided by complex implicit rules should remind us that Locke attacked, not only the theory of innate knowledge, but also the formalist doctrine according to which the foundation of all necessary truth is the possibility of formal derivation from such highest principles or 'maxims' as the law of non-contradiction, the general condition of all thought and all being whatsoever. On this doctrine, immediate insight into less general necessary relationships tends to become downgraded to the status of something more like 'intuition' in the modern sense, and to be interpreted as a sort of natural feeling for truth rather than explicit understanding. Full knowledge, it was held, would involve full logical derivation, the making explicit the reason why the truth in question is true. Locke's rival conception of the reasoning necessary for demonstrative knowledge presents it, as we have seen, as a series of particular or specific intuitions of relations between ideas. Some of such intuitions may be in accordance with the highest formal principles of logic, as (to take Locke's own examples) 'A circle is a circle' is in accordance with 'Whatever is, is'; or as 'Blue is not red' instantiates 'It is impossible for the same thing to be and not be'. Yet the possibility of that derivation is quite irrelevant to their evidence or certainty. The derivation of mathematical truths from a few maxims (or, as we might now add, from Russellian logic) can do nothing for our knowledge of them: 'how many truths are there about Numbers, which it is obvious to observe, that the Mind is perfectly acquainted with, and fully convinced of before it ever thought on these general Maxims?' Moreover, as we have seen, Locke went so far as to suggest that specific truths are more than chronologically first in our knowledge, for 'the general has more need to be let into [the] Mind by the particular, than the particular by the general'. Only because academics had become accustomed to employing certain very general propositions in disputation had it mistakenly come 'in time to be thought, that more particular Propositions have their Truth and Evidence from their Conformity to these more general ones'.¹⁴³ Locke's preference for the particular, echoing Gassendi's, if

faintly, is not entirely consonant with his general theory. For if intuition is infallible, it ought to follow that every intuition, whether very general or highly specific, is equally self-sufficient and independent of every other intuition.

'Intuitions' are not infallible, however, and the truth as to the value of principles seems to lie somewhere between Locke and his opponents. Locke is surely not wrong to deride that 'Principle of *Principles*, That Principles must not be questioned',¹⁴⁴ and the point can arguably be extended to logic itself. The principles of syllogism, for example, have no epistemological priority over particular arguments in accordance with those principles. Moreover any constructed formal system, however mathematically adequate or elegant, requires interpretation if it is to be of relevance to actual reasoning; and reasonable interpretation cannot neglect particular consequences.

Another consideration in favour of the independent force of examples can be drawn from the status of a *a priori* argument from analogy. Such argument cannot be explained as explicit deduction, but stands as an irreducible and logically formless type of *a priori* argument. It is not that such arguments lack all shape, but their cogency is not a function of their shape. They have the shape, '*a is f; therefore b is f*', and the appeal is to the judgement that *b* resembles *a* in such a way that it would be *inconsistent* to judge both that *a is f* and that *b is not f*. The inconsistency in question is not formal inconsistency, although it may be tempting to argue otherwise, i.e. to claim that *a* enters the argument simply as a reminder or illustration of the principle that whatever is *g is f*, so that the real form of the argument is '*b is g; all gs (such as a) are f; therefore b is f*'. Yet, first, there may be no such distinct predicate as '*g*' (e.g. if *f* is indefinable) and, second, the rewriting fails to capture the point of the original argument, which lies in a direct appeal to the analogy between two particulars.

On the other hand, even if particular *a priori* judgements have independent cogency, that cogency is not absolute. If particular intuitions seem to clash, we do characteristically look for general principles to make the situation clear. The general principle is then in some sense the explanation of more specific truths. We can conclude that general principles and specific judgements prop one another up, or can be employed to correct one another. Such a model is consonant with the view of understanding as the apprehension of connections, of significant resemblances and differences, or of a pattern in things: an apprehension which may be more or less partial but never, perhaps, absolutely comprehensive or exhaustive.

If all this comes near the truth, then the perception of necessity is something like seeing, but is more like ordinary cases of the intelligent apprehension of the significance of what we see, such as the realization how this (actual or depicted) gear works. More particularly, it is a certain kind of exercise of intelligent understanding in the context of language and linguistic techniques. 'Evidence', or the appearance of necessity, is simply the perspicuity which is a condition of primary understanding or knowledge. Not that we may not sometimes recognize necessary truth without achieving anything much worth calling 'understanding',

e.g. when the truth is as trivial or 'analytic' as 'All brothers are male'. Such understanding is little more than linguistic understanding, and the principle seems arbitrary rather than intelligible. In another kind of case we may arrive at an *a priori* conclusion by a mechanical procedure, as in arithmetic or formal logic, but here the procedure itself has at some point required intelligent understanding.¹⁴⁵ There may or may not be such a thing as innate mathematical ability, but doing mathematics makes some demands on general intelligence and the same intelligence as is required for the employment of mathematical symbolism is required for its acquisition. Moreover, even the notion of a specifically mathematical ability is not plausibly cashed out in terms of the possession of semantic principles. Yet mathematics supplies only one of the more obvious examples of how rationality and linguistic competence are intimately entangled. It is no doubt the failure to appreciate the implications of their interpenetration which is the cause of false theories of necessity and of *a priori* knowledge.

Conclusions of Volume I

More than the physical break between volumes makes this an appropriate place to take brief stock of what Locke's arguments have taught us. Roughly, the present volume has dealt with his general theory of knowledge and belief, particular and universal, and of the faculties which give rise to them, while the second volume will follow his account of their various objects, real and notional. The division is not a neat one, since epistemology and ontology do not so easily come apart: each shapes the other. Nevertheless, despite detours and foreshadowings, the argument up to this point has been dominated by a single theme: the question of the contribution of consciousness to thought and knowledge in their various forms. That question includes the question of the forms that consciousness itself can take, and of their relation to the epistemically primitive form of consciousness we call sensation.

From one point of view this emphasis on subjectivity or consciousness was hardly avoidable, but has simply been dictated by the presuppositions of Locke's argument: in effect, the presuppositions of his time. It was universally assumed that the primitive bearer of intentionality, that which in the first instance represents or has reference to objects, is conscious thought. As such, consciousness is the source of meaning, the intermediary between language and reality. For all but a few, in Locke's time as before, the structure of thought is the source of the structure of language. It was also generally assumed that our cognitive faculties are exercised in conscious quasi-perceptual or judgemental acts, and, in particular, that the crucial ingredient of knowledge or science lies in the conscious grasp of what is clear and evident and therefore immune to sceptical doubt. These ideas were not by any means new in the seventeenth century, as it has sometimes very oddly been supposed, but originated in ancient philosophy. They set the terms within which Locke constructed an epistemology fit for his undogmatic version of the 'new philosophy'. For this purpose, too, whether directly or indirectly, he drew on, and developed, ancient lines of argument. Like Epicureans and Stoics, he defended the authority and immediate evidence of the senses against both scepticism and intellectualism. He also developed an empiricist, imagist theory of abstraction and *a priori* knowledge

against dogmatic claims to have achieved a universal science of nature by penetrating to the essences of things.

The main, unfashionable philosophical thesis of the present volume has been that Locke's epistemology contains truth and can be of value to us, not *despite* being a theory about consciousness, but *became* it is a theory about consciousness. Not that one could possibly endorse Locke's conception of thoughts as composed of 'ideas', subjects and predicates lying before the mind and joined by acts of mental affirmation. Yet the possibility of predication does derive from the conditions and content of sensory experience and sensory knowledge. Primitive predication depends on our pre-linguistic powers to discriminate objects and be struck by resemblances between them, faculties exercised at the level of sensory consciousness, without which knowledge would be impossible. That topic, however, has been little more than broached in the arguments above, and will be explored in Volume II in the context of the question of how far our discriminations depend on us, and how far on reality. What has been considered more thoroughly in the present volume is the role of consciousness in knowledge, the role Locke marked by the notions of 'evidence' and the 'natural light'.

It has been argued at length above that modern analyses or explanations of knowledge, understanding and the like have been crippled by neglect of the characteristic 'perspicuity' of cognition. This neglect has been due in part to the influence of behaviourism and the analogy of the computer, which have given rise in many quarters to a dogged determination to leave subjectivity out of all account. Yet it has also been due in part to a reasonable rejection of the ancient idea, so popular in the seventeenth century, of a secure criterion of truth, the operation of an infallible knowledge-delivering faculty. Locke certainly did treat 'evidence' as such a criterion, and made it the foundation of an exclusive distinction between knowledge and belief. Yet it has been argued above that perspicuity can and should be distinguished from infallibility, and that in essence Locke's contrast between the perspicuous, self-confirming 'light of nature' and the essentially suspect 'light from heaven' can after all withstand criticism. The senses are not infallible, but perceptual knowledge is a paradigm (one might reasonably say, *the* paradigm) of knowledge just because of the way in which, in perceptual knowledge, there is no mystery to us as to why we believe what we believe. That is so, not in virtue of our having second-order beliefs about the origin of our belief (since those too must be perspicuous rather than blank hunches), but because in perception the source and object of belief is presented to consciousness. Perceptual belief is the natural acceptance of what is given in sensation, and as such it needs no credentials. There are, it is true, less perspicuous forms of knowledge (which Locke would have called 'belief' or 'opinion'), as when we remember something without remembering how we learnt it. We may even acquire knowledge subliminally, unconsciously. Yet these cases constitute secondary knowledge, parasitic on primary knowledge.

Although this argument endorsed something like Locke's conception of authoritative and evident perception (stripped of claims to infallibility), it did not endorse his account of its content, which his theory of simple ideas as the signs of their regular causes reduced to very little: in effect, to there being powers of things 'outside' us to cause sensations in us. That reduction was consonant with Locke's scepticism about essences, but it hardly does justice to the rich apprehension of the world which is immediately afforded by the senses, an integrated apprehension of things in space. Nevertheless a view in some respects structurally like Locke's was defended, drawing a firmer distinction than is now fashionable between the level of experience and the level of theory. Perceptual knowledge is foundational, although perceptual beliefs are defeasible.

If perceptual knowledge possesses, in such perspicuity, something like the 'evidence' taken by Locke to be an essential feature of knowledge, it may not be clear how the same perspicuity could pertain to *A priori* knowledge. In other words, how can *a priori* understanding be like sight? Locke took the analogy almost literally, but the imagist theory of universal knowledge which he developed fails to explain how modality (what is possible or necessary) is supposed to be apprehended by means of a single quasi-perceptual 'intuition'. Nominalist forms of the theory, like those of Hobbes and Hume, are no more successful, but point towards the modern linguistic theory of necessity which is neither imagist nor intuitionist. Yet, although it seems right that *a priori* reasoning is impossible without language, the linguistic theory fails to do justice to the natural and inescapable analogy with sight. No holders of the linguistic theory are vigilant enough to prevent themselves from ever talking of their 'seeing' when one thing follows from another, or that something must be so.

The issue is closely related to the question of how far, and how, intelligence is involved in the acquisition and possession of language. The linguistic theory fails to distinguish between blank linguistic intuitions and rational understanding; between recognition that a sequence is syntactically impermissible, or is gibberish, and recognition that, and why, what is being said does not make rational sense. We recognize logical or semantic impossibility as we recognize the impossibility of an Escher staircase as depicted. Similarly when we 'see' that, if A is larger than B, and B than C, then A is larger than C, we also see why that must be so. The grounding of our belief (which is not a storable reason) is immediately perspicuous to us. Although no more infallible than the evidence of the senses, such 'evidence' resists the traditional sceptical and anti-intuitionist reduction to mere subjective certainty. To *recognize* that *s* is *p* is not at all the same thing as to swear with conviction that *s* is *p*, or to be, in Locke's phrase, 'sure because one is sure'. *Recognition* and *understanding* are both categories which (*pace* Wittgenstein and others) as essentially involve consciousness as they are essential for an adequate characterization of consciousness, at any rate ordinary human consciousness. Here, of course, as with knowledge in general, it is necessary to distinguish between primary and secondary recognition and understanding.

Locke's epistemology raises many other rewarding issues, some of which have been discussed, whether rewardingly or not, above. But the dominant theme of the relationship of consciousness to knowledge (unless what has been argued is utterly remote from the truth) supplies perhaps sufficient illustration of that peculiar and deep feature of philosophical discourse and inquiry: it is possible to lose sight of what was once more clearly seen, and it is possible to discard truth together with the error with which it came entangled. More, it is almost impossible not to do these things. That is why it is possible for philosophy to renew itself, now as in the seventeenth century, by learning from the past.

Notes

Note

References are primarily by page number to the editions listed under ‘Bibliography’, but in order to facilitate reference to some multiply edited and collected primary sources (or to make active reference unnecessary for those sufficiently familiar with the texts) passages are in some cases also identified by short title and/or chapter, section etc. In the case of references to Locke’s *Essay*, the first line (only) of the passage is given after the page, thus: ‘Locke 1975:104, 13 (II.i.1)’, i.e. page 104, line 13. When a reference is to the last work referred to, the page number (or volume and page number, or, in the case of the *Essay*, page and line number) is given without the name of the author or date of the edition.

Part I: Ideas (pp. 11–77)

I Introduction to Part I

- 1 Locke 1958:200f: ‘the law of nature stands or falls with human nature as it now is’.
- 2 Cf. Locke 1975:516,28 (II.xi.16): ‘in moral discourses...[we] mean nothing by *Man*, but a corporeal rational Creature.... For were there a Monkey, or any other Creature to be found, that had the use of Reason,... he would no doubt be subject to Law.’
- 3 Locke 1958:148f; 1990:22 (cf. 50f, 75–8, 82).
- 4 Cf. Epicurus, cited in Long and Sedley 1987:I 87, II 91f: ‘First...we must grasp the things which underlie words, so that we may have them as a reference point...and not have everything indiscriminated for ourselves as we attempt infinite chains of proofs, or have words which are empty.’ Also Diogenes Laertius’ *Lives of the Philosophers* (of which Locke owned two copies) 10.33, cited in Long and Sedley 1987:I 87f, II 92f: ‘Preconception, they [the Epicureans] say, is as it were a perception, or correct opinion, or conception, or universal “stored notion” (i.e. memory) of that which has frequently become evident externally: e.g. “Such and such a kind of thing is a man.”’ Cf. Gassendi 1658:III 8f, and Locke 1958:148f: ‘the mind cannot discourse or reason without some truth that is given and

- perceived'. Contrast Locke 1975:595,3 (IV.vii.8), where the principle that 'all Reasonings are *ex praecognita, et praeconcessis*' is rejected.
- 5 Gassendi 1981:61.
 - 6 For a relevant comparison of Gassendi and Mersenne see James 1987.
 - 7 For charges of ambiguity (even of unintelligible or gross ambiguity) see e.g. O'Connor 1967:33–8; Hacking 1975:27f; Ryle 1968:17–20.
 - 8 But contrast O'Connor 1967:37f.
 - 9 Locke 1975:104,13 (II.i.1).
 - 10 159,21 (II.xi.9).
 - 11 150,6 (II.x.2, second edition).
 - 12 104ff (II.i.1–8); 117,10 (II.i.22); 155ff (II.xi); 163ff (II.xii).
 - 13 Cf. 120,3 (II.ii.2); 163,29 (II.xii.10). The theme is explored in Stewart 1979:53–82 and 1980:25–62. Cf. Ayers 1986:6–22.
 - 14 Locke 1975:43,17 (I.i.2). 'Physical' here of course means 'natural', and is not opposed to 'spiritual'.
 - 15 540,22 (IV.iii.6).
 - 16 576,12 (IV.v.6).

2

Ideas and compositionality in traditional logic

- 17 Aristotle 1984:I 26 (*De Interpretation [De Int.]* 17a 11–15).
- 18 For remarks about 'being said of and 'existing in' see I 3f (*Categories [Cat.]*. 1a 16–1b 9).
- 19 I.25 (*De Int.* 16a 4–16).
- 20 Fell 1673:2: 'Words are indicative signs of concepts and representative signs of things' (*voces sunt signa manifestativa conceptuum, suppositiva rerum*).
- 21 Hobbes 1839–45:I 35 (*De Corpore [De Corp.]* I.iii.7): 'A true proposition is that, whose predicate contains, or comprehends its subject, or whose predicate is the name of every thing, of which the subject is the name.'
- 22 For a discussion likely to have been familiar to Locke, see Smiglecki 1638:103–8.
- 23 Descartes 1964–76:X 363f (*Rules for the Direction of the Mind [Rules]* II).
- 24 VII 37 and 43f (*Meditations [Med.]* III).
- 25 VII 56–62 (*Med.* IV).
- 26 VIIIA 22–5 (*Principles of Philosophy [Princ.]* I 47–9).
- 27 Arnauld and Nicole 1965:33: '*le corps des preceptes...appartient [d'Aristote]*'.
- 28 Cf. Reid 1785:19: 'even a proposition may be simply apprehended without forming any judgment of its truth or falsehood: For it is one thing to conceive the meaning of a proposition; it is another thing to judge it to be true or false.' A traditionally more favoured way of dealing with the neutral consideration of a proposition was to suppose that it is the *sentence* which is really being considered. Cf. Nuchelmans 1980:104f.
- 29 Arnauld and Lancelot 1660:107–9.
- 30 Locke 1990:20 (*Draft A 9*).
- 31 Locke 1975:85,1 (I.iv.1).

- 32 405,15 (III.ii.1). Compare his promise at the end of the previous chapter to consider 'to what it is that Names, in the use of Language, are immediately applied'.
- 33 471,20 (III.vii.1). For a different view, a strong nominalism according to which affirmation is not an act of the mind combining ideas, but the combination of names itself, cf. Hobbes 1839–45:IV 23 (*Human Nature [Hum. Nat.]* v.9). Hobbes explicitly dismissed as absurd the view that *ens de ente praedicatur* at I 59 (*De Corp.* I.v.5).
- 34 Locke 1975:472,4 (III.vii.2); 472,10 (III.vii.2). Cf. Arnauld and Nicole 1965:104: 'Since all that takes place in the mind can be reduced to conceiving, judging, reasoning and ordering... words serve to signify (*marquer*) any of these activities.' The four activities correspond to the four parts of logic. Only in the first case do words stand for ideas.
- 35 Locke 1975:47,32 (I.i.8).
- 36 Locke 1976–88:IV 601f.
- 37 Locke 1990:15 (*Draft A 5*).
- 38 4 (*Draft A 1*).
- 39 8 (*Draft A 2*).
- 40 18 (*Draft A 8*). In another passage Locke employs two notions of simplicity, the new and the old, side by side: 47 (*Draft A 7*): 'And thus much for the Ideas of things simple or compounded... and the words by which men express them wherein I have been something the longer because not having met with it any where I thought soe new a notion... ought to be made plain. And this may be call'd the knowledg of simple termes and their ideas.'
- 41 Locke 1975:295,12 (II.xxiii.1).
- 42 305,5 (II.xxiii.14).

3

Ideas and epistemology before Locke

- 43 Descartes 1964–76:VI 85 (*Dioptrics [Diopt.]* I).
- 44 VI 112–14 (*Dioptr.* IV).
- 45 415–17 (*Rules* XII).
- 46 Hobbes 1839–45:I 389–91 (*De Corp.* IV.xxv.2). The introduction to this account indicates Hobbes' awareness of the peculiar nature of the problem of intentionality: 'Of all the phenomena or appearances which are near us, the most admirable is apparition itself...; namely, that some natural bodies have in themselves the patterns (*exemplaria*) almost of all things, and others of none at all.'
- 47 Hobbes 1839–45:I 396 (*De Corp.* IV.xxv.7): original, '*languescens vel debilitata sensio*'. Cf. IV 9 (*Hum. Nat.* iii.1): '*imagination being... conception remaining, and by little and little decaying from and after the act of sense*'.
- 48 Hobbes 1651:27 (*Leviathan [Lev.]* xxxiv); 1839–45:IV 426f.
- 49 Digby 1645:2f. Glanvill, for example, described Descartes and Digby as 'those two *Grand Sages*, then whom it may be the Sun never saw a more learned pair' (Glanvill 1661:33).
- 50 Cf. Boyle 1667, unpaginated 'Proœmial Discourse' (1979:8). Suarez argued explicitly that quantity and quality are distinguished from the substance in which

they inhere by a distinction *'ex natura rei'*, which is less than a fully real distinction but more than a mere distinction of reason, existing *'ante operationem intellectus'*. It is called a 'modal' distinction, in that *inherence* is a mode of quantity joining it to its substance (Suarez 1866:255–7). Descartes also allowed a difference between modal and conceptual distinctions, but on different grounds. Cf. Descartes 1964–76:VIII A 29f (*Princ.* I 61f) and VII 435 (*Sixth Reply*).

- 51 Hobbes 1839–45:I 102–5 (*De Corp.* II.viii.3).
 52 Boyle 1667:7 (1979:20f).
 53 The term 'accident' (as Boyle himself remarked) was ambiguous in Aristotelian philosophy. In the context of the doctrine of categories, whatever falls outside the category of substance is an accident (so all adjectival predicates predicate accidents.) In the context of the doctrine of 'predicates', however, 'accident' is opposed to 'difference' and 'property' as well as to 'species' and 'genus' (so only some adjectival predicates predicate accidents). Moreover, the predicables were applied to non-substantial 'species': in effect, to anything capable of definition. Criticism of 'real accidents' generally took 'accidents' to be opposed to 'substance', but also to be mutable and inessential attributes, opposed to 'properties'. For the most part, the distinction between senses of 'accident' seems to have been treated as irrelevant to the argument.
 54 'Real' as opposed to 'nominal', 'simple' as opposed to a composite of substance and accident, of which more below.
 55 Descartes 1964–76:VII 30–2 (*Med.* II). For the full argument to the geometrical essence of body, see VIII A 46 (*Princ.* II.11), which leads into the demonstration of the laws of motion.
 56 Descartes 1964–76:X 420f.
 57 Cf. Gassendi 1658:I 79–86; II 5ff; Lucretius, cited in Long and Sedley 1987:I 79, II 84 (*De Rerum Naturae* IV 499ff).
 58 Cf. Milton 1991.

4

Simple and complex ideas

- 59 Locke 1975:119,5 (II.ii.1).
 60 Aristotle 1984:I 4 (*Cat.* 1b25–2a10); Cf. I 33 (*De Int.* 20b12–21a18); II 1626 (*Metaphysics [Metaph.]* 10 1029b22–1030a18).
 61 Arnauld and Nicole 1965:46–9. The passage will be discussed in Volume II, chapter 5.
 62 Locke 1975:563,27 (IV.iv.3).
 63 564,4 (IV.iv.4).
 64 Cf. 375,8 (II.xxxi.2); 388,15–390 (II.xxxii. 14–16). Cf. also Gassendi 1658:11 9. The form of argument adopted here and its classification by Epicurus and others is discussed in Long and Sedley 1987:91–6. For the doctrine of indicative signs, see chapter 3, above.
 65 Descartes 1964–76:VIII A 34 (*Princ.* I 70).
 66 VII 43f (*Med.* III); 206f; 231–5.
 67 VIII A 32–6 (*Princ.* I 66–71). Cf. Arnauld and Nicole 1965:71f.
 68 Locke 1975:119,18 (II.ii.1).

- 69 Leibniz 1875–90:IV 422f (‘Reflections On Knowledge, Truth and Ideas’).
 70 Locke 1975:388,23 (II.xxxii.14); 373,3 (II.xxx.2).
 71 132,14 (II.viii.1–7).
 72 201,7 (II.xv.9). Cf. 167,10 (II.xiii.2f).
 73 234,19 (II.xxi.3).

5

Ideas as images

- 74 Gassendi 1658:II 440–6.
 75 Hobbes 1839–45:I 20 (*De Corp.* I.ii.9).
 76 See notes 14 and 15, above.
 77 Locke 1975:297,24 (II.xxiii.5).
 78 542,34 (IV.iii.6). Cf. Locke 1823:IV 463f.
 79 47,28 (I.i.8).
 80 Gassendi 1658:I 92. Cf. I 33, and contrast Descartes 1964–76:VII 181.
 81 Locke 1975:134,11 (II.viii.7).
 82 372,20 (II.xxx.2).
 83 152,3 (II.x.5).
 84 155,13 (II.x.10). Another echo of Gassendi, who had brought the case of a dog attuning its barks to the sound of a trumpet against Descartes’ claim that animals are mere machines. Cf. Descartes 1964–76:VII 270.
 85 Locke 1975:117,13 (II.i.22).
 86 98,23 (I.iv.20).
 87 347,18 (II.xxvii.27).
 88 396,6 (II.xxxiii.6).
 89 698,22 (IV.xix.3).
 90 562,24 (IV.iv.1).
 91 396,24; 400,7 (II.xxxiii.6 and 17).
 92 118,6 (II.i.24).
 93 368,26 (II.xxix.13f). Locke was probably aiming to improve on Gassendi’s objection to *Meditations* VI (Descartes 1964–70:VII 329ff): ‘although you perceive that the word “chiliagon” signifies a figure with a thousand angles, that is just the meaning of the term, and it does not follow that you *understand* the thousand angles of the figure any better than you *imagine* them’.
 94 Locke 1975:206,15 (II.xvi.5f).
 95 Hobbes 1651:14 (*Lev.* iv).
 96 Locke 1975:216,29 (II.xvii.13).
 97 363,18 (II.xxix.2).
 98 119,18 (II.ii.1).
 99 105,23 (II.i.4). For the useful discussion of problems facing Locke with respect to the notion of reflection, see Kulstad 1984.
 100 Locke 1823:IV 19.
 101 Locke 1975:172,37 (II.xiii.13).
 102 131,8 (II.vii.7).
 103 205,8 (II.xvi.1).
 104 550,12 (IV.iii.19).

- 105 Descartes 1964–76:X 368, a passage which also speaks of the fluctuating testimony of the senses.
 106 Locke 1975:633,13 (IV.xi.6).
 107 680,31 (IV.xvii.8).

6

Ideas as intentional acts and ideas as intentional objects

- 108 Descartes 1964–76:VII 40–5 (*Med.* III).
 109 VII 8 (*Med.* Preface).
 110 VII 92–4 and 102–5.
 111 Malebranche 1980:217–20, 230–41 and 612–33 (*Search* III.ii.1, 6 and 7; *Elucidation* X).
 112 Arnauld 1683:13–58.
 113 46–9.
 114 Yolton 1984:88–115. Cf. Gibson 1958:19f.
 115 Reid 1785:143–56 (*Essays on the Intellectual Powers of Man [Essays]* II.ix).
 116 Arnauld and Nicole 1965:39.
 117 Locke 1975:680,32 (IV.xvii.8).
 118 229,25 (II.xx.2).
 119 596,7 (IV.vii.9).
 120 134,20 (II.viii.8).
 121 138,5 (II.viii.17).
 122 137,10 (II.viii.15).
 123 138,11 (II.viii.18).

7

Ideas as natural signs

- 124 Locke 1975:525,1 (IV.i.1). Descartes' and Arnauld's views, and Yolton's interpretation of Locke, are discussed in chapter 6, above.
 125 Locke 1823:IX 211–55.
 126 Arnauld 1683:25f.
 127 Locke 1975:720,30 (IV.xxi.4).
 128 Cf. 574,6 (IV.v.2).
 129 121,22 (II.iii.1).
 130 105,6 (II.i.3).
 131 630,25 and 635,12 (IV.xi.2 and 9).
 132 Cf. Locke 1823:IX 218; 1975:229,22 (II.xx.2).
 133 Locke 1975:137,30 (II.viii.16); cf. 136,30 (II.viii.13).
 134 375,20 (II.xxxi.1); cf. 137,14 (II.viii.15).
 135 146,8 (II.ix.8).
 136 127,12 (II.v).
 137 Cf. Lievers, forthcoming, which establishes the importance of this context for both Molyneux and Locke.
 138 Descartes 1964–76:VI 137f (*Diopt.* VI 13).
 139 Locke 1975:145,26 (II.ix.8).

Ideas: a summary

- 140 525,5 (IV.i.2).
 141 Arnauld 1683:37.

Reflections on the structure of thought

- 142 This (not, of course, the only modern) approach to meaning is classically presented by Wittgenstein, who was much inclined to target for criticism the traditional assumption that what gives sentences their life lies in consciousness. In rejecting that assumption, he seems to have found the whole conception of thoughts as inner states and processes problematic, such states being no more than constructed shadows of sentences (a view discussed in chapter 22, below). But a conception of 'inner' states other than the traditional one has become influential in present-day philosophy: a 'functionalist' can hold that an inner physiological state of a kind systematically related to the sentence uttered is always a determinant of the meaning of any utterance. Moreover, Paul Grice's approach to meaning (Grice 1957, etc.) presupposes that such inner states as intentions and beliefs are prior to language. Grice's critics, on the other hand, complain that he offers no account of the content of mental states which is language-independent (cf. Schiffer 1987), a complaint consonant with the view that 'the only proper method for analysing thought consists in the analysis of language' (Dummett 1978:458). Thus modern attitudes towards Locke might be more complex than any simple contrast between Locke and Wittgenstein would suggest: for example, a Gricean could endorse *both* Locke's view that thought is prior to language *and* Wittgenstein's critique of the assumption that meaning is determined by ideas capable of being brought 'before the mind' as elements of the content of consciousness.
- 143 Locke 1975:574,7 (IV.v.2). Cf. Hobbes 1651:15 (*Lev.* I.iv).
- 144 Cf. Stalnacker 1984:4: 'Representational mental states should be understood primarily in terms of the role that they play in the characterization and explanation of action.... Linguistic action, according to this picture, has no special status.'
- 145 Cf. Fodor 1975; Field 1978.
- 146 Locke 1975:635,18 (IV.xi.9)

*Part II:**Knowledge and belief (pp. 79–150)**Introduction to Part II*

- 1 Aristotle 1985:I 165f (*Posterior Analytics* 99b15–100b17).
 2 Cf. Long and Sedley 1987:I 238–59.

- 3 Cf. James 1987, to which this resumé of Mersenne's epistemology is indebted.
- 4 Descartes 1964–76:IXB 27 (*Princ.* I 58).
- 5 Arnauld and Nicole 1965:291f.
- 6 Gassendi 1981:60–4.
- 7 30f.
- 8 61. In Arnauld and Nicole 1965:316f the principle is rejected that 'all the certainty and evidence of propositions comes either immediately or mediately from the senses': 'Our belief in the truth of the axiom "The whole is greater than the part" will only be probable if it rests on no more than our having observed that a man is larger than his head,...the sky than a star.... The certainty of this axiom...depends solely on what the clear and distinct ideas we have of a whole and of a part clearly contain.... All that the various observations we have made...can have done is to have given us occasion to attend to [these] ideas.' Gassendi himself, it should be said, included 'the whole is greater than the part' among useful maxims or necessary propositions which are both the most evident and the most general and which are such that an understanding of their meaning is a sufficient condition for assent. Yet 'necessary' here seems to have boiled down to nothing more than 'always true'. Thus 'Every man is rich' and 'No man is rich' both come out as 'impossible' rather than 'contingent', while 'Some man is rich' and 'Some man is not rich' are both 'necessary'. 'Coriscus is playing' is contingent, but 'The sun is many times greater than the earth' is necessary, although evident only through reasoning based on experience (Gassendi 1981:29f). The principle that maxims need only be understood in order to command assent seems to have been interpreted by Gassendi in the following way: anyone having ordinary extensive experience of particular wholes and parts who is then presented with, and comes to understand, the abstract general maxim, will assent to it. The maxim draws together our universal experience.
- 9 Gassendi 1981:64. Cf. 36f.
- 10 Hobbes 1839–45:IV 9 (*Hum. Nat.* v.9). Cf. I 30 (*De Corp.* I.v.2) etc. He rejects the notion of one thing's being predicated of another thing (*ens de ente praedicatur*) as absurd at I 59 (I.v.5).
- 11 I 35 (*De Corp.* I.iii.7); IV 28 (*Hum. Nat.* vi.3).
- 12 These two paragraphs are based on Hobbes 1651:18–32 (*Lev.* I.v–vii) and Hobbes 1839–45:IV 26–30 (*Hum. Nat.* vi).
- 13 Locke 1975:653,24 (IV.xiv.3f). 'Belief', 'opinion' and 'judgement' are not always equivalent terms in the *Essay*, however. At 528,31 (IV.i.9) Locke follows Hobbes' usage of 'belief': 'Bare Belief...relies on the Testimony of another.' 'Judgment', moreover, unlike 'belief' or 'opinion', is often used for the faculty as well as for its exercise. At 156,27 (II.xi.2) it is used for a different faculty, that of avoiding a certain kind of fallacy in judging. Note that, as generally used for the cognitive act or state, Locke's 'judgement' and 'belief' include both Hobbes' 'judgement' (based on past experience) and his 'belief' or 'faith' (based on testimony). Gibson 1968: 121 states that 'judgement' suffers from an 'unfortunate ambiguity' in the *Essay*, meaning not only *opinion*, but, in general, any 'act of thought by which an affirmation or denial is made,' including those which constitute knowledge. This claim seems to be without the least foundation.
- 14 Locke 1975:524,5 (IV.i.1).
- 15 655,17 (IV.xv.3).

- 16 654,9 (IV.xv.1); 656,19 (IV.xv.5).
- 17 525,16 (IV.i.3). Cf. 543,31 (IV.iii.8f).
- 18 527,30 (IV.i.8f).
- 19 530,21 (IV.ii).
- 20 As in IV.iv, although the distinction between modes and relations is not important for the account of knowledge.

11

The degrees of knowledge and the role of method

- 21 Locke 1975:531,2 (IV.ii.1). Cf. 618,20 (IV.ix.3).
- 22 532,8 (IV.ii.2).
- 23 685,28 (IV.xvii.18).
- 24 683,25 (IV.xvii.14). But at 543,21 (IV.iii.6), he takes it that angels use intermediate ideas, but easily find and remember them. For a modern analogue of this question, cf. Ayer 1958:85f.
- 25 533,15 (IV.ii.6f).
- 26 Cf. Aaron 1963:220–7. Aaron’s view is criticized in Schankula 1980.
- 27 Descartes 1964–76:X 369 (*Rules III*).
- 28 Cf. Spencer 1628:8.
- 29 Descartes 1964–76:X 368 (*Rules III*).
- 30 Arnauld and Nicole 1965:235f.
- 31 Descartes 1964–76:X 363 (*Rules III*).
- 32 X 430–7 (*Rules XIII*). Much of this passage was available more or less verbatim in the Port Royal *Logic* (Arnauld and Nicole 1965:300–6).
- 33 Descartes 1964–76:VI 64f (*Discourse VI*).
- 34 Locke 1975:670,5 (IV.xvii.4).
- 35 673,25 (IV.xvii.4).
- 36 Gassendi 1981:40 (III.i).
- 37 Locke 1975:681,11 (IV.xvii.8). Cf. 675,10 (IV.xvii.4).
- 38 668,27 (IV.xvii.2f).
- 39 680,9 (IV.xvii.7). The phrase is quoted from Hooker, but cf. the use of ‘*auxiliis veris intellectus*’ at Bacon 1878:167.
- 40 Locke 1975:682,1 (IV.xvii.9f).
- 41 597,22 (IV.vii.10).
- 42 603,17 (IV.vii.11). The argument is essentially that of Gassendi 1981:61, but moderated and retailed for a less radical conclusion than Gassendi’s. Gassendi differed from his contemporary and friend Mersenne, who argued in *La Verité des Sciences* that the proposition that the body is greater than its fingers derives certainty from the maxim that the whole is greater than its parts (cf. note 8, above). Locke was in all probability consciously differing from both writers on this issue.
- 43 Locke 1975:599,23 (IV.vii.10). Cf. Bacon 1878:191: ‘The sciences we now have are just a sort of expositions of what has already been discovered, not ways of making discoveries (*modi inveniendi*) or specifications for new works (*designationes novorum operum*).’

- 44 Locke 1975:647,29 (IV.xii.12). Cf. 641,36 (IV.xii.4); 109,10 (II.i.10), etc.
 45 644,3 (IV.xii.9–12).
 46 643,10 (IV.xii.7). Cf. 649,30 (IV.xii.15) and 682,22 (IV.xvii.11).
 47 A programme advocated chiefly in Book III of the *Essay*. See (in particular) chapter 32, below, and Volume II, chapter 6.
 48 537,15 (IV.ii.14).
 49 Gassendi 1981:31 (II.xiv).
 50 Locke 1975:537,4 (IV.ii.14).
 51 662,10 (IV.xvi.6).
 52 663,3 (IV.xvi.9).
 53 Cf. 701,25 (IV.xix.10): ‘For [knowledge and belief] are two ways, whereby Truth comes into the Mind, wholly distinct, so that one is not the other.’
 54 Hume 1968:188–93.
 55 Hence the problem of the ‘reality’ of knowledge, as raised at Locke 1975:562,19 (IV.iv.1–3), and the point of arguments at 633,13 (IV.xi.6) and at 550,10 (IV.iii.19).
 56 706,13 (IV.xx.1). Cf. Descartes 1964–76:X 368 (*Rules* III): ‘Because [intuition] is simpler, it is more certain than deduction, though deduction...is not something a man can perform wrongly’ (and VII 146, etc.).

12

Other divisions of knowledge

- 57 Locke 1975:527,32 (IV.i.8).
 58 The most influential text here probably being Wittgenstein 1958.
 59 Locke 1975:56,23 (I.ii.17–23).
 60 528,6 (IV.i.8).
 61 Cf. 97,22 (I.iv.20): ‘For what is not either actually in view, or in the memory, is in the mind no way at all, and is all one as if it never had been there.’ Locke made this point right at the beginning of his attack on innate principles, at 49,30 (I.ii.5): ‘It seeming to me near a Contradiction, to say, that there are Truths imprinted on the Soul, which it perceives or understands not.’
 62 528,30 and footnote (IV.1.9). This is ‘belief’ in the narrow sense explained in chapter 10 above, although the argument also seems to require that ‘belief’ and ‘opinion’ are equivalent.
 63 530,4 (IV.i.9).
 64 Compare Descartes 1964–76:VII 69–71 (*Med.* V) and 140f (*Second Reply*). Descartes argued that knowledge dependent on remembered clear and distinct perceptions also depends (if it is to count as *scientia*, free from ‘metaphysical’ doubt) on the principle that God, not being a deceiver, would not constitute us in such a way that propositions remembered as evident might turn out to be false. Hence the atheist is incapable of *scientia*. An unfortunate feature of Locke’s fourth edition argument is what seems to be an attempt to find a non-metaphysical rival to this principle, in the principle ‘What *Ideas* once agreed will always agree’, which he proposed as an ‘intermediate idea’ requisite if remembered demonstrative knowledge is to count as knowledge. Thus the ‘Agreement of the two *Ideas* join’d in that Proposition is perceived, but it is by the intervention of other *Ideas* than

those which at first produced that Perception.’ Moreover, upon ‘this ground it is, that particular demonstrations in Mathematics afford general knowledge’. This running together of his explanation of the universality and eternity of the eternal truths (cf. Locke 1975:638,30) with his account of remembered knowledge is an awkward red herring. It would seem not only to make, for example, the grounds of remembered mathematical knowledge non-mathematical (although the knowledge is still, allegedly, demonstrative), but to import a kind of universalizing principle into every demonstration as a premise. Nevertheless a more sensible view of memory as retained knowledge exists alongside this ill-conceived flourish.

- 65 Russell 1959:114f; 1956:157–87. Ryle 1958:272–9.
 66 Locke 1975:525,19 (IV.i.3). Cf. 543,33 (IV.iii.7), where the terminology and order are unimportantly different.
 67 Cf. 364,11 (II.xxix.5).
 68 611,1 (IV.viii.3).
 69 609,12 (IV.viii.2).
 70 527,20 (IV.i.7).
 71 Cf. 596,1 (IV.vii.9).
 72 In the work of Noam Chomsky, and that of his critics.
 73 Locke 1975:612,3 (IV.viii.3).
 74 613,9 (IV.viii.5).
 75 Hobbes 1839–45:1 62 (*De Corp.* I.v.10).
 76 Locke 1975:614,30 (IV.viii.8).
 77 Further ground for this conclusion is afforded by the discussion of the issue in Locke 1990:51–61 (*Draft A* 27–31).
 78 Cf. Locke 1975:644,16 (IV.xii.9):

Experience here must teach me, what Reason cannot: and ’tis by trying alone, that I can certainly know...whether that yellow, heavy, fusible Body, I call Gold, be malleable, or no; which Experience (which way ever it prove, in that particular Body, I examine) makes me not certain that it is so, in all, or any other yellow, heavy, fusible Bodies, but that which I have tried.

The thought here is awkward, seeming first to imply that I *can* know from experiment that (all) gold is malleable, and then to state that I can’t know it. Yolton 1970:113 quotes another passage, Locke 1975:665,21 (IV.xvi.12), which seems to imply that we can have knowledge of such universal coexistential propositions as ‘the Load-stone draws iron’, but again we should take the point to concern only cases of which we have had experience. Other passages are unequivocal, e.g. 546, 31 (IV.iii.14): ‘For this *co-existence* can be no farther known, than it is perceived... either in particular Subjects by the observation of our Senses, or in general, by the necessary *connexion* of the *Ideas* themselves.’

- 79 543,10 (IV.iii.6). Cf. 618,20 (IV.ix.3).

- 80 Locke 1975:653,24 (IV.xiv.4).
- 81 654,4 (IV.xv.1).
- 82 657,26 (IV.xvi.1).
- 83 663,24 (IV.xvi.9).
- 84 Locke 1990:62 (*Draft A* 33). Elsewhere in *Draft A* ‘faith’ is used synonymously with ‘assent’ and ‘opinion’. For an account of late Scholastic discussions of the distinction between ‘apprehensive’ and ‘judicative’ mental propositions, see Nuchelmans 1980:74–113. Locke’s early proposal seems close to that ascribed to Gregory of Rimini: ‘if a proposition is merely entertained, without any assent or dissent, it is not a mental proposition in the strict sense, but a mental image of a spoken or written proposition and thus a conventional sign’.
- 85 Donald Davidson claims that it is impossible to have a belief without the concept of a belief (and so without language), in effect extending the argument from ‘apprehensive’ to ‘judicative’ prepositional thought. But his premise begs the question: ‘Someone cannot have a belief unless he understands the possibility of being mistaken’ (Davidson 1984:170).
- 86 Locke 1975:706,16 (IV.xx.1).
- 87 Cf. 718,12 (IV.xx.16), which contains the phrase ‘perceives the greater Probability’; and 669,3 (IV.xvii.2), according to which the mind may come to perceive or see a ‘probable connexion’ between ideas.
- 88 711,16 and 715,16 (IV.xx.7 and 12).
- 89 652,1 and 653,1 (IV.xiv.1 and 3).
- 90 Cf. 655,15 (IV.xv.3) and, especially, 685,20 (IV.xvii.17): ‘*Judgement*, is the thinking or taking two *Ideas* to agree, or disagree, by the intervention of one or more *Ideas*, whose certain Agreement, or Disagreement with them it does not perceive, but hath observed to be frequent and usual.’
- 91 Passmore 1986:25.
- 92 Locke 1975:650,28 (IV.xiii.2).
- 93 717,10 (IV.xx.16).
- 94 717,20 (IV.xx.16).
- 95 716,12 (IV.xx.15).
- 96 Cf. Passmore 1986:32: ‘It would certainly be quite ridiculous to say we ought not to perceive the landscape until we have examined it more closely.’
- 97 Locke 1975:715,34 (IV.xx.14).
- 98 659,32 (IV.xvi.4).
- 99 715,31 (IV.xx.14).
- 100 716,15 (IV.xx.15).
- 101 706,25 (IV.xx.1).
- 102 708,15 (IV.xx.3). Cf. Locke 1823:IV 154, and Dunn 1969:231ff.
- 103 Locke 1975:709,29 (IV.xx.5): ‘there is a difference of degrees in Men’s Understandings, Apprehensions and Reasonings’.
- 104 But perhaps not, in Locke’s view, always blameworthy: cf. 396–401 (II.xxxiii.7–18).
- 105 713,12 (IV.xx.10).
- 106 712,21; 714,9; 714,36; 718,22; 719,20 (IV.xx.9, 11, 12, 17 and 18).
- 107 717,6 (IV.xx. 15): ‘that a Man should afford his Assent to that side, on which the less Probability appears to him, seems...as impossible, as it is to believe the same thing probable and improbable at the same time’.

- 108 715,14 (IV.xx.12). My understanding of this passage differs from Passmore's.
- 109 In fact Locke's summary, at 717,2 (IV.xx.15), fits a paradigm case of someone who will not use available proofs, rather than someone who uses 'wrong measures': 'In... less clear Cases...it is in a Man's Power to... content himself with the Proofs he has, if they favour the Opinion which suits with his Inclination, or Interest.'
- 110 Locke nowhere, I think, states that we *inevitably* respond to probabilities, although he does imply that in certain cases we do so respond to *apparent* probabilities. Cf. 716,16 (IV.xx.15).
- 111 Locke 1975:696 (IV.xix.1).
- 112 Contrast Passmore 1986:39 and 42. As Locke explains them, the enthusiast's beliefs differ from the lover's in that their content is not intrinsically connected to the passions which maintain the self-delusion.
- 113 Locke 1975:698,22; 699,27; 703,7 (IV.xix.3, 7 and 11).
- 114 699,32 (IV.xix.7).
- 115 700,4 (IV.xix.8).
- 116 702,4 (IV.xix.10).
- 117 Locke 1936:120 (19 February 1682): 'that which makes all these pretences to supernaturall illumination...farther to be suspected to be merely the effect and operation of the phansy is that all the preparation and ways used to dispose the minde to these illuminations...are such as are apt to disturb and deprese the rationall power of the minde, but to advance and set on worke the phansy'.

14

The grounds of probability

- 118 Locke 1975:656,1 (IV.xv.4).
- 119 661,10 (IV.xvi.5). The distinction corresponds to the ancient distinction between empirical and indicative signs, discussed in chapter 3, above.
- 120 657,1 (IV.xv.6).
- 121 661,30 (IV.xvi.6).
- 122 662,15 (IV.xvi.7).
- 123 662,25 (IV.xvi.8).
- 124 663,7 (IV.xvi.9).
- 125 Russell 1959:66f.
- 126 Locke 1975:661,33 (IV.xvi.6).
- 127 666,35 (IV.xvi.12).
- 128 Cf. Bacon 1878:181–4, 195–9; 1858–61:IV 40–3,50.
- 129 Locke 1975:648,4 (IV.xii.13).
- 130 Bacon 1878:255; 1858–61:IV 71.
- 131 Locke 1975:647 (IV.xii.12).
- 132 547,30 (IV.iii.16).
- 133 Bacon 1878:278; 1858–61:83.
- 134 Locke 1975:134,28 (II.viii.9).
- 135 136,1 (II.viii.11).
- 136 308,14 (II.xxiii.23); 545,23 and 558,29 (IV.iii.13 and 28). These passages are discussed in Volume II.

- 137 10,2 (Epistle). Rogers 1978 argues persuasively that philosophical additions to the second edition of *Principia* were influenced by Locke.
- 138 Hume 1957:109–31.
- 139 Locke 1975:559,25 (IV.iii.29). The point of interpretation, currently controversial, will be argued more fully in Volume II.
- 140 Locke 1989:246 (*Education* 192).
- 141 Locke 1823:X 255.
- 142 IX 256.
- 143 IX 263.
- 144 IX 257–60.
- 145 Locke 1975:667,8 (IV.xvi.13).
- 146 Descartes 1964–76:VII 84.
- 147 Locke 1975:691,21 and 692,7 (IV.xviii.5).
- 148 694,24 (IV.xviii.8).
- 149 667f (IV.xvi.14). The same line of argument at 695,27 (IV.xviii.27) opens with the statement, ‘Whatever GOD hath revealed, is certainly true; no Doubt can be made of it.’ This might suggest a confusion over an ambiguity of scope, but Locke was imputing just such confusion to those who made dogmatic appeals to revelation. The entirely conventional principle that revelation leaves no room for doubt was listed by Gassendi among the maxims of logic without Locke’s sceptical proviso (Gassendi 1981:36 and 64).
- 150 Cf. Popkin 1983.
- 151 Suggested in conversation by Thomas Lennon. Cf. Malebranche 1980:573; Berkeley 1964:214. But Berkeley used the expression in the context of discussion of the doctrine that we see all things in God, while Locke’s use of a physiological explanation of enthusiasm is reminiscent of Malebranche’s own discussions of imaginative delusion. Yet that leaves the question whether Locke’s immediate motive for adding the chapter to the fourth edition, apparently so long after the actual civic harm supposed to have been caused by English enthusiasts, was simply that of filling an obvious lacuna in his argument.
- 152 Locke 1975:388,25 (II.xxxii.14). Cf. 372,25 (II.xxx.2); 375,9 (II.xxxi.2); 564,2 (IV.iv.4).
- 153 631,26 (IV.xi.3).
- 154 692,33 (IV.xviii.5).
- 155 Ashcraft 1969.
- 156 Locke 1975:698,25 (IV.xix.4).
- 157 702,33 (IV.xix.11).
- 158 700,27 (IV.xix.9).
- 159 531,10 (IV.ii.1).
- 160 700,25 (IV.xix.9). Cf. Descartes 1964–76:VII 191f.

15

Reflections on the definition of knowledge

- 161 Locke, as an imagist, in effect questions whether the distinction between literal and metaphorical seeing is as sharp as this way of putting it would suggest, but we may abstract from that complication for the time being.

- 162 Locke 1975:655,23 (IV.xv.3).
 163 Cf. Gettier 1963. We all owe a debt to this spectacularly effective squib, but it is more difficult, and probably pointless, to try retrospectively to assign a source, stimulus or irritant cause for each of the thoughts that follow.
 164 Cf. Russell 1959:132.
 165 The proponent of this line of argument will no doubt have equally optimistic plans for dealing with intentional content and truth.
 166 Nozick 1981:167–96.
 167 Peacocke 1986:127–70.
 168 Cf. 154f; Nozick 1981:190
 169 The general point is an important one, on which Peacocke seems to falter. On ‘abduction’ he writes:

First, an explanatory suggestion is proposed which would explain certain phenomena; and second, the hypothesis is accepted as true on the basis of further experiment, evidence or reasoning. The method consists not just of thinking up a simple hypothesis to explain the data: it also comprises a policy of not accepting the hypothesis as true unless one has a sufficient range of evidence which, in the circumstances, ensures that if the hypothesis were not true, one would not believe it.... Thus for instance it is by now the case that if DNA were not to have the shape of a double helix, we would not believe it to do so. Until this counterfactual became true, the hypothesis about its shape was not knowledge.

(Peacocke 1986:140)

Maybe so, but if the suggestion is that the Nozickian hypothetical supplies an independent explanatory criterion for estimating the point at which supported belief topples over knowledge, it needs to be said that the hypothetical judgement itself, so far from being independently based on the psychology of believers, simply follows an assessment or evaluation of the evidence for the hypothesis.

- 170 As has often been pointed out, a minor qualification is necessary with respect to this feature: roughly, some of the subject’s reasons for a belief may be true and some false, and yet the belief still constitutes knowledge, provided that the true grounds or premises by themselves (i) constitute decisive reason for the belief and (ii) would have influenced the subject to believe even if he or she had known that the false premises were false.
- 171 An awkward issue should here be mentioned, i.e. the question of knowledge and beliefs about our present conscious states. Such beliefs (*pace* some present-day philosophers) are not plausibly taken to derive from their objects through the deliverances of a self-scanning faculty with its own mechanism. Having a pain or visual sensation is too closely tied up with knowing or being aware that one has it for that model to be adequate. The issue will receive some discussion in chapter 22, below. For the present, a note only.
- 172 Locke 1975:529,13 (IV.i.9). Cf. the discussion of habitual knowledge in chapter 12, above.

- 173 As at 697,15, 18 and 21 (IV.xix.1).
 174 703,31 (IV.xix.13).

16

Belief and rationality

- 175 Descartes 1964–76:VII 59 (*Med.* IV).
 176 Locke 1975:697,4 (IV.xix.1).
 177 Cf. Wright 1983. Textual evidence for this interpretation, which represents a turning-point in our understanding of Hume, is comprehensively cited in G.Strawson 1989.

*Part III:**Perceptual knowledge (pp. 151–217)*

18

The authority and limits of 'sensitive knowledge'

- 1 Locke 1975:630,27 and 631,1 (IV.xi.2). The whole argument is strikingly (and, one supposes, significantly) like the view attributed to the Stoic Chrysippus by the author of *De Placitis Philosophorum*:

An impression is an affection occurring in the soul, which reveals itself and its cause. Thus when through sight we observe something white, the affection is what is engendered in the soul through vision: and it is this affection which enables us to say that there is a white object which activates us.

(Long and Sedley 1987:I 237)

Locke possessed this work, both in Greek and in a Latin translation, in Stephanus' 1572 edition of Plutarch.

- 2 Cf. Descartes 1964–76:VII 80–90 (*Med.* VI); VIII A 34–6 (*Princ.* I.70f); Arnauld and Nicole 1965:71f.
 3 Locke 1975:631,17 (IV.xi.3).
 4 537,15 (IV.ii.14).
 5 634,26 (IV.xi.8). Emphasis on 'certain' added.
 6 630,30 (IV.xi.2).
 7 Arnauld and Nicole 1965:337f. Cf. 293.
 8 Locke 1975:713,12 (IV.xx). The victim has accepted 'as a foundation of Reasoning, That he must believe his Reason (for so Men improperly call Arguments drawn from their Principles) against their [*sic*] Senses.' It seems virtually certain that Locke had the passage from the Port Royal *Logic* in mind.
 9 The arguments summarized are those of 632–5 (IV.xi.4–8).
 10 631,14 (IV.xi.3). The translation of Lucretius' *De Rerum Natura* IV, lines 474f, 478f and 499, is from Long and Sedley 1987:78f.
 11 Cf. Locke 1975:527,9 (IV.i.7).

- 12 140,7 (II.viii.22).
- 13 Locke 1823:IV 7; cf. 29 and 33.
- 14 Locke 1976–88:VI 630.
- 15 Locke 1975:295,13 (II.xxiii.1) The meaning of this passage will be the subject of extended argument in Volume II.
- 16 635,18 (IV.xi.9).
- 17 298,19 (II.xxiii.6).
- 18 233,10 (II.xxi.1).
- 19 324,7 (II.xxvi.1).
- 20 325,20 (II.xxvi.2).
- 21 294,18 (II.xxii.11).
- 22 Cf. 44,4 (I.i.2). One might take the meaning of ‘historical’ to overlap partially the meaning of the currently fashionable ‘naturalistic’ (if the latter has any clear meaning).
- 23 235,5 (II.xxi.4).

19

Does perceptual knowledge have independent authority?

- 24 Locke 1975:706,14 (IV.xx.1).
- 25 139,21 (II.viii.21). Here Locke himself identifies ‘the Sensation of Heat and Cold’ with the increase or diminution of motion, in line with his common use of the term ‘sensation’ for the physiological process in the sense organ. For the general thought, compare Gassendi 1658:II.7: ‘All appearances or sensations are true.... The diversity of appearances misleads some [in this matter].... For they suppose that of discrepant or contradictory appearances one has to be true, and the other, from its opposition to truth, to be false. But that is actually a crude assumption, by people who do not consider the nature of things.’ Cf. Long and Sedley 1987:I 81, and chapter 3, above. Another of Locke’s arguments reminiscent of this Epicurean doctrine is his treatment of the supposition that physiological differences could cause systematic and undetectable differences in the way people see colours.
- 26 Gassendi 1658:III 7. Cf. 1981:12.
- 27 Or we might say that I see a face in my shirt, as we see faces in the fire. Cf. chapter 9, above and 31, below.

20

Does perceptual knowledge have a firm boundary?

- 28 Kant 1963:41 (*Critique*, Introduction i).
- 29 Peacocke 1983:19.
- 30 Rorty 1980:106. Cf. Churchland 1979:2: ‘we are forced to acknowledge the demise of the familiar distinction between theoretical beliefs and perceptual beliefs. The class of perceptual beliefs must now be counted as a subclass of the class of theoretical beliefs: roughly as those singular theoretical beliefs acquired as spontaneous non-inferential responses to sensory states of the perceiver. As with singular theoretical judgements generally then, the adequacy of our perceptual judgements is in part a matter of the adequacy of the background theory (conceptual

framework) in whose terms they happen to be framed.’ A line of argument refreshingly contrary to this dogma is advanced in Fodor 1983, to the effect that sensory ‘input-systems’ exist as ‘modules’, so that sensory information, the end-product of a sensory process, is in general unaffected by wider beliefs. Unfortunately Fodor takes *perception* to involve belief, and perceptual belief to be a product of a ‘central system’ responding to sensory information in the light of wider beliefs. This looks like a version of the old idea that perception involves ‘conceptualization’. Yet having appropriately produced sensory information is perceiving, and to have perceptual beliefs is simply to accept what the senses deliver.

- 31 Quine 1965:40–6.
 32 Peacocke 1983:7, 100 and 89.

21

The scope of perceptual knowledge

- 33 Locke 1975:138,8 (II.viii.18).
 34 200,15 (II.xv.8): ‘Where and when are Questions belonging to all finite Existences’.
 35 Cf. Berkeley 1964:I 15 (*Philosophical Commentaries [Phil. Comm.]* 80); II 239 (*Three Dialogues between Hylas and Philonous [Three Dial.]* III); I 256.
 36 It may be worth pointing out that ‘smell’ is systematically ambiguous: the smell of an object may be thought of as an attribute like its colour or size, or as an effect of that attribute which spreads over an area commonly centred on the object (but not if there is a strong wind).
 37 Locke 1975:122–7 (II.iv). Cf. Descartes 1964–76: VIII A 42 (*Princ.* II.4).
 38 Cf. Locke 1975:545,16 (IV.iii.13).
 39 Hume 1968:230f (*A Treatise concerning Human Nature [Treat.]* I.iv.4).
 40 Cf. Malebranche 1923:300.
 41 Reid 1785:235–48.
 42 Cf. Rorty 1980:96f. According to Rorty’s not wholly unreasonable estimate (although things may be changing), the problem of consciousness has become in recent literature ‘primarily... a problem about pains’ as blank ‘raw feels’ rather than a problem about intentional states. Rorty himself rejects the view that every mental state is intentional, giving pain as the prime counter-example (22).
 43 Cf. Sacks 1989:42–52.
 44 Berkeley 1964:I 211 (*An Essay towards a New Theory of Vision [New Theory]* sect. 98f).

22

Two modern approaches to sensation

- 45 Nagel 1979:165–7.

- 46 An extended discussion of this sort of issue from the functionalist point of view occurs at Dennett 1981:149–73. At different points in his argument Dennett seems inclined towards each of the responses here outlined.
- 47 The natural order of explanation is reversed in the functionalist's attributing the intentional content of our sensory states to the content of what their occurrence inclines us to believe or say (rather than attributing the content of those declarations or beliefs to the content of antecedent sensory states). This helps to explain the strong and interesting alliance which exists in writers like Quine, Rorty and Churchland between behaviouristic approaches to intentionality and anti-foundationism (although there are other reasons for it). Both behaviourism (including functionalism) and anti-foundationism have difficulties in allowing that the intrinsic content of sensation available to phenomenological reflection sets a boundary round perceptual belief and knowledge.
- 48 An extreme case of such anthropomorphism, reported by Nagel, is the apparently serious application of the categories of belief and desire to a thermostat, which was said to desire the surrounding air to remain at 65° F and from time to time to believe that its temperature has fallen below that point, the desire and belief jointly issuing in action. Less extreme functionalists face the curious task of explaining why thermostats *don't* have minds.
- 49 Cf. Wittgenstein 1958:135, 162f; Anscombe 1957; Davidson 1980:3–19; 207–25. Wittgenstein's notion of first-person 'expressions' of sensation is discussed at Malcolm 1963:106–17).
- 50 The assumption that the category of 'behaviour' is unproblematic has long been criticized on the ground that the notion of 'behaviour' itself presupposes that of purpose. In view of this cogent point, it is significant that a version of functionalism has arisen according to which the primitive purposiveness required by any 'externalist' explanation of the intentionality of experience and thought is supplied by biological need: ideas, beliefs and intentions are such (and are intentional) 'because of what they are, given the context of their [evolutionary] history, *supposed* to do and of how they are supposed to do it' (Millican 1984:93). A false belief is in this like a malfunctioning heart or the germination of a seed in the wrong season, and indeed the flow of adrenalin is said to be even more narrowly like a belief in being an indicative 'intentional icon' which is, however, also imperative (116ff). The position generates some paradoxes peculiar to it, but deserves fuller consideration than can be given here.
- 51 Cf. Nuchelmans 1986 and 1988.
- 52 Locke 1975:472,6 (III.vii.2).
- 53 473,10 (III.vii.5).

23

Private language and secondary qualities

- 54 Locke 1975:375,24 (II.xxxi.1). The passage is discussed in chapter 7, above.
- 55 389,4 (II.xxxii.14).
- 56 389,9 (II.xxxii.15).
- 57 Cf. Hacker 1976.
- 58 Locke 1975:390,6 (II.xxxii.16). Cf. note 25, above.

- 59 Kripke 1972:353f.
 60 Infra-red light is, of course, otherwise perceptible, as radiant heat.
 61 Descartes, of course, did argue that hardness is a sense-relative attribute. The argument is not a strong one: cf. Descartes 1964–76:VIII A 42.
 62 Cf. Grice 1962:134f.
 63 This is too simple and in another context would need qualification in the light of the persuasive sceptical argument of Hardin 1988 (to which I owe the expression ‘electronic structure’) that there are so many variables in the physics and physiology of colour vision that it is impossible to give a non-arbitrary account of ‘standard’ conditions under which objects will appear their ‘real’ hues. Nevertheless the full complexity of the facts he adduces seems irrelevant to the general points being made here. All that is assumed here is what Hardin would presumably concede, that the pre-linguistic discrimination of things by their colours is *sufficiently* founded on the structure of the objects seen as coloured to be useful; and, correspondingly, that the general agreement at a rough and ready, everyday level which makes colour predication possible is *largely* founded on that structure.

Part IV:

Particulars, universals and intuitive knowledge (pp. 219–300)

25

Locke’s arguments on space and time in context

- 1 Locke 1975:200,15 (II.xv.8).
 2 Aristotle 1984:354–69 (208a26–217b28).
 3 369–74 (217b29–220b33).
 4 Cf. Descartes 1964–76:VIII A 42–52 (*Princ.* II 5–20); II 479; V 267.
 5 VIII A 52 (*Princ.* II. 21).
 6 VIII A 26f (*Princ.* I 55–7).
 7 Hobbes 1839–45:I 93f and 99 (*De Corp.* II.vii.2 and 12).
 8 I 109 (*De Corp.* II.viii.10).
 9 I 410–25 (*De Corp.* IV.xxvi.1–4).
 10 I 102 (*De Corp.* II.viii.1).
 11 I 106–8 (*De Corp.* II.viii.5–7).
 12 I 94–5 (*De Corp.* II.vii.3).
 13 Gassendi 1658:I 179–228.
 14 More 1743:I 75–125; 1662:19–22.
 15 Locke 1990:45 (*Draft A* 27).
 16 Locke 1990:223–53 and 259f (*Draft B* 101–30 and 141–4).
 17 Cf. Descartes 1964–76:64–5 (*Med.* V).
 18 Locke 1975:167,24 (II.xiii.4).
 19 Cf. Berkeley 1964:I 16 (*Phil. Comm.* 87): ‘The idea of an inch length not one determin’d idea.’

- 20 The example is perhaps a good one to counter the notion that a simple recognitional skill involves the application of a ‘concept’. Possessing the concept *40cm long* is irrelevant to the capacity to judge 40cm lengths by eye. The concept of the measurement does not bestow the capacity, while someone gifted with the capacity may lack any system of measurement (the capacity may be the capacity to judge by eye whether something is as long as any given medium-sized object, e.g. an object 40cm long). The ability to carry out the measurement is not itself a recognitional capacity, although it involves the primitive capacity to appreciate equality. Cf. chapter 20, above.
- 21 Cf. Berkeley 1964:I 204–6 (*New Theory* sects 80–7); II 45 (*A Treatise concerning the Principles of Human Knowledge [Princ.]* I 11), II 60 (*Princ.* I 47), II 97–102 (*Princ.* I 123–32). The arguments are discussed in Ayers 1982:61–5.
- 22 Not that Berkeley would have been likely to have held that, in the Müller-Lyer case, there are visual lines of different lengths. By analogy with his treatment of the moon on the horizon, he would have held that the visual lines are of the *same* length (composed of an equal number of *minima*) but that the mind *judges* them to be different. In effect, he would have denied the occurrence of the explicandum, a *visual* illusion. Cf. Berkeley 1964:I 196–203 (*New Theory* sects 67–78).
- 23 The example is J.L.Austin’s.
- 24 Leibniz 1981:147.
- 25 Arnauld and Nicole 1965:294f; Malebranche 1980:25–32 (*The Search after Truth [Search]* I.vi); Bayle 1710:IV 3079 (‘Zeno’).
- 26 Cf. Locke 1975:187–91 (II.xiv. 17–22). The *Draft B* discussion of time had been largely aimed at breaking the traditional link with motion in order to make room for the primitive internal clock from which the idea of duration is derived. See especially Locke 1990:223–38 (*Draft B* 101–12).
- 27 Berkeley 1964:83 (*Princ.* I 98).
- 28 Leibniz 1981:152.
- 29 Kant 1963:244–7 (*Critique of Pure Reason* B274–9).
- 30 See Journal entries for 27 March and 20 June 1676; 16 September 1677; 20 and 24 January 1678 (all in Locke 1936:77–105).
- 31 Locke 1975:179,19 (II.xiii.26).
- 32 180,7 (ibid.). The whole argument summarized here extends over 171–80 (II.xiii. 11–26).
- 33 173,14 (II.xiii.14).
- 34 198,1 (II.xv.4).
- 35 171,4 (II.xiii.10).
- 36 Aaron 1963:160f. The present account is largely consonant with the suggestions made in Gibson 1958:245–54 (and attacked by Aaron), although Gibson was unaware of *Draft C*, and did not discuss place.
- 37 *Draft C* xviii, especially 10 and 12.
- 38 i.e., at Locke 1975:171,4 (II.xiii.10).
- 39 As at 199,19 (II.xv.6).
- 40 198,35 (II.xv.5). Cf. 171,16: ‘the undistinguishable *Inane* of infinite Space’.
- 41 Newton 1962–6:I 10–12. The other thought-experiment, presumably in hostile allusion to Aristotle’s bucket of water, involves a spinning bucket whose absolute motion would be revealed through the water’s rising up its sides.

26

Thought about particulars

- 42 Locke 1975:329,4 (II.xxvii.2). Cf. the reference to God in this section with 306,29 (II.xxiii.19f). In the context of these passages, and the chapters on space, the difficulty raised at 542,24 (IV.iii.6) of reconciling ‘Existence to any thing that hath no Extension at all’ (a difficulty purportedly inherent in the notion of immaterial spirit) must be presumed to be the difficulty of understanding how something might have location without extension. More’s account of spirits had at any rate avoided that difficulty.
- 43 159,14 (II.xi.9).
- 44 Nor, of course, is there any sign of the Hobbesian view that the absolute place of an object is available to us demonstratively, with the object itself, a view incompatible with the quasi-Newtonian conception of absolute place introduced (if the present interpretation is right) into the first edition of the *Essay*. For our idea of the absolute place of an object is for Locke irremediably obscure, and certainly not available in sense-perception. So his notion of absolute place could offer no way out of the circle by providing ostensibly identifiable ‘known points’.
- 45 Locke 1975:199,1 (II.xv.5). Cf. 201,2 (II.xv.8) etc.

27

Locke’s theory of universal knowledge in context

- 46 Arnauld and Nicole 1965:47–59. Strictly speaking Arnauld differentiates the distinction between universal and particular ideas from the distinction between ideas which represent many things and ideas which represent one thing, since ‘some man’ is taken to stand for a particular idea, representing many things.
- 47 Descartes 1964–76:VIII A 27–32 (*Princ.* I 58–64).
- 48 Arnauld 1683:104.
- 49 Cf. Long and Sedley 1987:179–83; 237f. The explanation of universal truths as mere hypothetical was used by Suarez as an argument for denying that eternal essences need to have been created (Suarez 1866:II 231).
- 50 Long and Sedley 1987:180.
- 51 Gassendi 1981:6.
- 52 10f.
- 53 Descartes 1964–76:VII 319–22.
- 54 Hobbes 1839–45:I 29–37 and 44–55 (*De Corp.* I.iii.1–8 and iv); IV 14–24; 28 (*Hum. Nat.* iv, v 1–10, vi 3). Cf. Hobbes 1651:8–17 (*Lev.* I.iiif).
- 55 Hobbes 1651:14 (*Lev.* I.iv).
- 56 Hobbes 1839–45:I 20 (*De Corp.* I.ii.9). For the relation between science and prudence, see IV 21 and 28f (*Hum. Nat.* v 4 and vi 4), and Hobbes 1651:21f (*Lev.* I.v).
- 57 Locke 1975:159,10 (II.xi.9). The word ‘precise’ as it appears in this passage (at 159,18), might have been read as a synonym for ‘abstract’ (rather than for ‘exact’),

had it not been the case that elsewhere, as at 580,4 (IV.vi.4), it is transparently equivalent to 'exact', which itself appears at 580,13.

58 562,6 (IV.iii.31).

59 Cf. Leibniz 1981:49: 'The senses...never give us anything but instances, that is particular or singular truths.' Leibniz, of course, concludes that something more than sense or imagination is required for universal knowledge.

60 Locke 1975:681,1 (IV.xvii.8).

61 As Locke put it in *Draft C* (quoted by Aaron 1963:65), the mind 'makes use but of one idea to contemplate all existing of that kind; whereby that one idea becomes as it were a representative of all particulars that agree with it'.

62 Locke 1975:638,30 (IV.xii.14).

63 Cf. note 49, above and Hobbes 1839–45:I 38 (*De Corp.* I.iii.10): 'those propositions only are *necessary*, which are of sempiternal truth, that is, true at all times...; for it will be eternally true, *if man, then living creature*'. Yet it is evidently important to Locke (and essential to the argument of *Essay* III.iii.19) that the hypothetical which explains why philosophers think of such truths as eternal relates not (as for Hobbes) directly to the world, but to our ideas:

being once made, about abstract *Ideas*, so as to be true, they will, whenever they can be supposed to be made again at any time past or to come, by a Mind having those *Ideas*, always actually be true....

[Hence] Propositions, concerning any abstract *Ideas*, that are once true, must needs be *eternal Verities*.

(Locke 1975:639,1)

64 Berkeley 1964:II 27–40 (*Princ.* Introduction).

65 Aaron 1963:193–219.

66 Locke 1975:172,4 and 173,2 (II.xiii.11 and 13). Contrast Berkeley's accusation that for Locke 'the mind can frame to itself by abstraction the idea...of motion exclusive of...extension' (Berkeley 1964:II 28).

67 Berkeley 1964:II 48 (*Princ.* I 17).

68 Locke 1975:131,10 (II.vi.7).

69 205,8 (II.xvi.1).

70 208,27 (II.xvi.7).

71 205,16 (II.xvi.3–5). Cf. 551,15 (IV.iii.19); 535,6 (IV.ii.10).

72 55,25 and 56,1 (I.ii.15f).

73 566,10 and 633,14 (IV.iv.8 and IV.xi.6); 567,6 (IV.iv.9). Cf. 550,14 (IV.iii.19): 'An Angle, Circle or Square drawn in Lines, lies open to the view and cannot be mistaken: It remains unchangeable, and may at leisure be considered.' I.e. the employment of the senses and sensible objects in demonstration lends a clarity and stability lacking when the objects of reasoning are ideas of the imagination.

74 Locke 1990:22 (*Draft A* 11). Cf. 50f (*Draft A* 27) where his earlier knowledge-empiricism seems to have made way for the possibility of *a priori* knowledge:

And indeed demonstrations are...as the word denotes the bare shewing of the things or proposing them to our senses or

understandings...certain knowledg or demonstration makes it self clearly appeare and be perceived by the things them selves put together before our senses or their clear distinct Ideas put together and as it were lyeing before us in view in our understandings.

In Locke 1823:IV 59 (*Stillingfleet*) we still find 'this showing or demonstration'.

- 75 Hobbes 1839–45:I 60 and 20 (*De Corp.* I.v.8 and ii.9).
 76 Locke 1975:414,5 (III.iii.11).
 77 575,11 (IV.v.2).
 78 579,20 (IV.vi.2).
 79 Cf. 159,7 (II.xi.9), where it is assumed that we form abstract 'standards' in order to make general names possible. Cf. Berkeley 1948:36 (*Princ.*, Introduction sect. 18).
 80 Locke 1975:159,35 (II.xi.10). Cf. 402,15 (III.i.3) etc.
 81 Cf. Hobbes 1651:14 (*Lev.* I.iv): 'without words, there is no possibility of reckoning of Numbers'.
 82 Locke 1975:575,1 (IV.v.4). Cf. 579,7.
 83 Cf. Hobbes 1839–45:I 62 (*De Corp.* I.v.10): 'For every proposition, universally true, is either a definition, or a part of a definition, or the evidence of it depends upon definitions.'
 84 Hume 1968:21 (*Treat.* I.i.7).
 85 19f.
 86 One of the most famous is a drawing of a square staircase which rises to the level of its own base.

28

Abstraction and the ideal of precision

- 87 A special context, however, might determine that the sentence should count as true, e.g. if A and B both satisfy some ordinary colour-predicate, or if they are sufficiently alike for some purpose in view.
 88 See below, chapter 30.
 89 Locke 1975:580,3 (IV.vi.4). Problems are supposed to arise over substances because another boundary than exact conformity with the representative idea 'can be supposed' mistakenly (i.e. a natural boundary). Cf. 470,9 (III.vi.50).
 90 122,8 (II.iv.2).
 91 224,21 and 33 (II.xviii.4 and 6).
 92 319,20 (II.xxv.1); 349,1 (II.xxviii.1). Cf. chapter 26, above.
 93 205,24 (II.xvi.3); 535,10 (IV.ii.10).
 94 At 365,5 (II.xxix.7) we are told that 'tis complex *Ideas* that are most liable to confusion'. At 479,35 (III.ix.9), simple ideas and ideas of substances are contrasted with ideas of mixed modes, as being less liable to confusion since their names are taught ostensively. Yet at 535,18 (IV.ii.11–13) a reason is given why we cannot nicely distinguish those simple ideas 'whose Modes and Differences are made, and counted by degrees, and not quantity': i.e. not knowing the corpuscular causes of those degrees, 'we have no certain Standard to measure them by'.
 95 Berkeley 1964:I 128 and 123f.

Intuition and innate knowledge

- 96 Cf. Fodor 1983.
- 97 Cf. Locke 1975:49,30 (I.ii.5): 'it seeming to me near a Contradiction to say that there are truths imprinted on the Soul, which it perceives or understands not; imprinting, if it signify any thing, being nothing else, but the making certain Truths to be perceived'. Cf. also 97,5 (I.iv.20): 'Whatever *Idea* was never perceived by the mind, was never in the mind.'
- 98 Descartes 1964–76:VIII 358.
- 99 Locke 1975:63,6 and 64,25 (I.iii.25–7). Such flourishes can be matched by, and were very probably modelled on, arguments of English *proponents* of innatism who wished to make it clear that their doctrine concerned dispositions or 'anticipations', not actual or conscious knowledge with understanding. Yolton quotes N.Culverwel, from a work of 1654: 'had you these connate *species* in the *Cradle?* and were they rock't asleep with you? or did you then meditate upon these principles, *Totum est majus parte, & Nihil potest esse & non esse simul?*' (Yolton 1956:42). Yolton analyses the long-running, largely theological debate in England to which Locke was a late contributor, arguing that a naïve theory of actual, luminous innate knowledge was replaced by a more sophisticated dispositional theory. Yet the evidence rather supports the conclusion that no one at all held a truly 'naïve' theory (for who could?). Some writers were prepared to stress that talk of 'imprinted' or 'engraved' truths, of 'seeds of light', 'heavenly beams' and the like is potentially misleading metaphor, while others took up a less cautious rhetorical stance. Yet all were thinking in terms of dispositions of whatever strength. Yolton also seems to assume that Locke's targets were English, but it is impossible that Locke should have been unaware of the implications of his attack for Cartesian philosophy (so frequently and explicitly criticized elsewhere in the *Essay*), or, indeed, of Descartes' reputation among many English writers.
- 100 Locke 1975:50,37 (I.ii.5).
- 101 58,4 (I.ii.18).
- 102 Leibniz 1981:76f.
- 103 Locke 1975:528,6 (IV.i.8).
- 104 It is astonishing how slow proponents of the theory were to take note of the point of this criticism: cf. James Lowde's inadequate response to Samuel Parker (quoted by Yolton 1956:53): 'it must be granted, that it is at least possible for God to imprint truths upon the minds of Men: and if so then it cannot be done any other way, than by making 'em thus plain and self evident'. The identification of innateness with quasi-logical self-evidence was pretty general, even if John Hartcliffe, publishing just after the *Essay*, spoke of innate inclinations which 'do not proceed from Reason, but from Nature', and which are 'antecedent to all Reason and Discourse' (Yolton 1956:48f). In general, Burthogge's hostile characterization of 1675 was clearly appropriate: 'in the *speculative* Understanding they have set up a habit which they call *Intelligence*; in the *Practical* another which is called *Synteresis*; in both a Constellation of *Principles*, shining with their own Light, and imparting it to others that want it' (Yolton 1956:46f). Locke's very satisfying position was that principles are evident to us in virtue of our general faculty

of reason, not in virtue of their innateness. As he commented in the margin of Thomas Burnet's *Third Remarks on the Essay*, 'this is an improper way of speaking to call a power principles' (Yolton 1956:56).

30

Locke on meaning and some modern criticisms

- 105 Locke 1975:409,27 (III.iii.2).
 106 407,3 (III.ii.4).
 107 407,9 (III.ii.5).
 108 407,12 (III.ii.5).
 109 Cf. Putnam 1975:II 139–52 and 215–90.
 110 Cf. Locke 1975:427,9 (III.iv.15); 486,32 (III.ix.18f); 503,25 (III.x.22). The linking of simple ideas with modes (presumably both simple and complex) in this respect at 407,12 (III.ii.5) is therefore atypical and, indeed, at odds with the immediately consequent reference to mixed modes in particular.
 111 Arnauld and Nicole 1965:83–5.
 112 Descartes 1964–76:VIII A 22 (*Princ.* I 45).
 113 Locke 1975:363,10 (II.xxix.2–5). The explanation of clarity (363,17) is reminiscent of Stoic accounts of cognitive or 'kataleptic' impressions, which are said to be clear, distinct and striking (*tranes, ektypos, enarges, plektike*). Cognitive impressions arise when there is no impediment to the senses (compare Locke's 'in a well-ordered sensation or Perception'). Cf. Long and Sedley 1987:I 242 and 246, II 244 and 250.
 114 Locke 1975:365,3 (II.xxix.7–12).
 115 509,3 (III.xi.1). Cf. 578,15 (IV.v.10) and 510,15 (III.xi.5), quoted below.
 116 Cf. 425,20 (III.iv.11). At 536,32 (IV.ii.13), Locke claimed, as the conclusion of a rather obscure argument over four sections, that '*Ideas of Colours, as we see in different kinds, as Blue and Red, are as capable of Demonstration, as Ideas of Number and Extension.*' But we do not need to adopt the strongly imagist conception of self-evidence and demonstration behind this claim in order to take the present point.
 117 649,5 (IV.xii.15).
 118 476,20 (III.ix.3).
 119 509,8 (III.xi.2).
 120 521,24 (III.xi.24); 492,4 (III.x.4). Cf. 407,30 (III.ii.7); 437,6 (III.x.15) etc. for this recurring theme.
 121 719,31 (IV.xx.18).
 122 101,8 and 25 (I.iv.23).
 123 Cf. 517,12 (III.xi.17); 504,3 (III.x.22) etc.
 124 Cf. 514,11 (III.xi.11).
 125 522,31 (III.xi.25). Cf. 479,13 (III.ix.13) etc.
 126 406,11 (III.ii.3).

31

Reflections on understanding and imagination

- 127 Locke 1975:406,29 (III.ii.4).
 128 Wittgenstein 1958:58 (they are, of course, the words of his translator, G.E.M.Anscombe).
 129 Russell 1956:200.
 130 208.
 131 227.
 132 Wittgenstein was very possibly influenced by a reading of *The Analysis of Mind* in arriving at his view that the content of consciousness is not significantly relevant to meaning or understanding.
 133 Anscombe 1957:42.
 134 49.
 135 Putnam 1975:3–7.
 136 Cf. Sacks 1989:130.

32

Necessity, reason and language

- 137 Descartes 1964–76:178f.
 138 Chomsky 1968:24. Contrast Chomsky’s sharply different statement on page 68 of the same work, that the ability to acquire language is a ‘species-specific capacity that is essentially independent of intelligence’.
 139 Katz 1966:1–6, 186–317.
 140 132–41.
 141 Wisdom 1957:249–52.
 142 Katz 1966:286f. Cf. 92f.
 143 Locke 1975:595,14; 603,20 and 29 (IV.vii.9 and 11).
 144 101,34 (I.iv.24).
 145 Katz himself explains the three variables, *input—language acquisition-device—output*, with the point that ‘a very intelligent person can obtain the solutions to certain mathematical problems (the output) given just the barest formulation of the problems (the input) whereas a very unintelligent person might have to be virtually told the solutions before he gets them’. But the assimilation of intelligence to a ‘device’ hardly helps his argument (Katz 1966:174f).

Bibliography

- Aaron, R.I. 1963: *John Locke* (Oxford)
- Alexander, P. 1985: *Ideas, Qualities and Corpuscles* (Cambridge)
- Anscombe, G.E.M. 1957: *Intention* (Oxford)
- Aristotle 1985: *The Complete Works of Aristotle: The Revised Oxford Translation*, ed. J.Barnes (Oxford)
- Arnauld, A. 1683: *Des Vrayes et des Fausses Idées* (Cologne)
- Arnauld, A. and C.Lancelot 1660: *Grammaire Générale et Raisonnée* (Paris) [Solar Press facsimile]
- Arnauld, A. and P.Nicole 1965: *La Logique ou l'Art de Penser*, ed. P.Clair and F.Girbal (Paris)
- Ashcroft, R. 1969: 'Faith and Knowledge in Locke's Philosophy', in *John Locke: Problems and Perspectives*, ed. J.Yolton (Cambridge)
- Ayer, A.J. 1958: *Language, Truth and Logic* (London)
- Ayers, M.R. 1982: 'Berkeley's Immaterialism and Kant's Transcendental Idealism', in *Idealism Past and Present*, ed. G.Vesey (Cambridge)
- 1986: 'Locke's Logical Atomism', in *Rationalism, Empiricism and Idealism*, ed. A.Kenny (Oxford)
- Bacon, F. 1858–61: *Works*, ed. J.Spedding, R.L.Ellis and D.D.Heath (London)
- 1878: *Novum Organum*, ed. T.Fowler (Oxford)
- Bayle, P. 1710: *An Historical and Critical Dictionary* (London)
- Berkeley, G. 1964: *The Works of George Berkeley*, ed. A.A.Luce and T.E.Jessop (London)
- Boyle, R. 1667: *The Origine of Formes and Qualities* (Oxford)
- 1744: *Works* (London)
- 1979: *Selected Philosophical Papers of Robert Boyle*, ed. M.A.Stewart (Manchester)
- Chomsky, N. 1968: *Language and Mind* (New York)
- Churchland, P.M. 1979: *Scientific Realism and the Plasticity of Mind* (Cambridge)
- Coke, Z. 1654: *The Art of Logick* (London) [Solar Press facsimile]
- Copleston, F. 1962: *A History of Philosophy* (New York)
- Davidson, D. 1980: *Essays on Actions and Events* (Oxford)
- 1984: *Truth and Interpretation* (Oxford)
- Dennett, D.C. 1981: *Brainstorms: Philosophical Essays on Mind and Psychology* (Brighton)
- Descartes, R. 1964–76: *Oeuvres de Descartes*, ed. C.Adam and P.Tannery (Paris)
- 1970: *Descartes: Philosophical Letters*, ed. A.Kenny (Oxford)
- 1985: *The Philosophical Writings of Descartes*, trans. J.Cottingham, R.Stoothoff and D.Murdoch (Cambridge)
- Digby, K. 1645: *Two Treatises: Of Bodies and Of Man's Soule* (London)
- Dummett, M. 1978: *Truth and Other Enigmas* (London)

- Duncan, W. 1748: *Elements of Logic* (London) [Solar Press facsimile]
- Dunn, J. 1969: *The Political Thought of John Locks* (Cambridge)
- Fell, J. 1673: *Grammatica Rationis: Sive Institutiones Logicae* (Oxford)
- Field, H. 1978: 'Mental Representation', *Erkenntnis* 13:9–61
- Fodor, G. 1975: *The Language of Thought* (Hassocks)
- 1983: *The Modularity of Mind* (Cambridge, Mass.)
- Gassendi, P. 1658: *Opera Omnia* (Lyons)
- 1981: *Institutio Logica 1658*, ed. H.Jones (Assen)
- Gettier, E. 1963: 'Is Justified True Belief Knowledge?', *Analysis* 23, 6:121–3
- Gibson, J. 1968: *Locke's Theory of Knowledge and its Historical Relations* (Cambridge)
- Glanvill, J. 1661: *The Vanity of Dogmatizing* (London)
- Grice, H.P. 1957: 'Meaning', *Philosophical Review* 66, 3:377–88
- 1962: 'Some Remarks about the Senses', in *Analytical Philosophy: Second Series*, ed. R.J.Butler (Oxford)
- Hacker, P.M.S. 1976: 'Locke and the Meaning of Colour Words', in *Impressions of Empiricism*, ed. G.Vesey (London)
- Hacking, I. 1975: *Why Does Language Matter to Philosophy?* (Cambridge)
- Hardin, C.L. 1988: *Colour for Philosophers* (Indianapolis)
- Hobbes, T. 1651: *Leviathan* (London) [Solar Press facsimile]
- 1839–45: *The English Works of Thomas Hobbes*, ed. W.Moles-worth (London)
- Hume, D. 1957: *Enquiries concerning the Human Understanding and the Principles of Morals*, ed. L.A.Selby-Bigge (Oxford)
- 1968: *A Treatise of Human Nature*, ed. L.A.Selby-Bigge (Oxford)
- James, S. 1987: 'Certain and Less Certain Knowledge', *Proceedings of the Aristotelian Society* LXXXVII: 227–42
- Kant, I. 1963: *The Critique of Pure Reason*, trans. N.Kemp Smith (London)
- Katz, J.J. 1966: *The Philosophy of Language* (New York)
- King, P. 1829: *The Life of John Locke* (London)
- Kripke, S. 1972: 'Naming and Necessity', in *Semantics of Natural Language*, ed. D.Davidson and G.Harman (Dordrecht)
- Kulstad, M. 1984: 'Locke on Consciousness and Reflection', *Studia Leibnitiana* XVI, 2: 143–67
- Leibniz, G.W. 1875–90: *Philosophischen Schriften*, ed. C.I.Gerhardt (Berlin)
- 1981: *New Essays on Human Understanding*, trans. P.Remnant and J.Bennett (Cambridge) [pagination follows edition of A.Robinet and H.Schepers, Berlin 1962]
- Lievers, M. forthcoming [1991]: 'The Molyneux Problem', *Journal of the History of Philosophy*
- Locke, J. 1823: *The Works of John Locke* (London)
- 1936: *An Early Draft of Locke's Essay*, ed. R.I.Aaron and J.Gibb (Oxford)
- 1958: *Essays on the Law of Nature*, ed. W. von Leyden (Oxford)
- 1963: *Two Treatises of Government*, ed. P.Laslett (London)
- 1975: *An Essay concerning Human Understanding*, ed. P.H.Nidditch (Oxford)
- 1976–88: *The Correspondence of John Locke*, ed. E.S.de Beer (Oxford)
- 1989: *Some Thoughts concerning Education*, ed. J.W.Yolton and J.S.Yolton (Cambridge)
- 1990: *Drafts for the Essay concerning Human Understanding and other Philosophical Writings*, Vol. I, ed. P.Nidditch and G.A.J.Rogers (Oxford)
- Long, A.A. and D.N.Sedley 1987: *The Hellenistic Philosophers* (Cambridge)

- McCracken, C.J. 1983: *Malebranche and British Philosophy* (Oxford)
- Mackie, J.L. 1976: *Problems from Locke* (Oxford)
- Malcolm, N. 1963: *Knowledge and Certainty* (Englewood Cliffs)
- Malebranche, N. 1923: *Dialogues on Metaphysics and on Religion*, trans. M.Ginsberg (London)
- 1980: *The Search after Truth*, trans. T.M.Lennon and P.J.Olscamp (Columbus, Ohio)
- Mellor, D.H. 1985: *Real Time* (Cambridge)
- Millican, R.G. 1984: *Language, Thought and Other Biological Categories* (Cambridge, Mass.)
- Milton, J.R. 1991: 'Locke and Gassendi', in *Studies in Seventeenth-Century Philosophy*, ed. M.A. Stewart (Oxford)
- More, H. 1662: *The Immortality of the Soule*, in *A Collection of Philosophical Writings* (London)
- 1743: *Divine Dialogues* (Glasgow)
- Nagel, T. 1979: *Mortal Questions* (Cambridge)
- Newton, I. 1962: *Philosophiae Naturalis Principia Mathematica*, trans. A.Motte (rev. F.Cajori) (Berkeley)
- Nozick, R. 1981: *Philosophical Explanations* (Oxford)
- Nuchelmans, G. 1980: *Late Scholastic and Humanist Theories of the Proposition* (Amsterdam)
- 1982: *Judgement and Proposition from Descartes to Kant* (Amsterdam)
- 1986: 'The Historical Background to Locke's Account of Particles', *Logique et Analyse* 113: 53–71
- 1988: 'The Distinction *Actus Exercitus/Actus Significatus* in Medieval Semantics', in *Meaning and Inference in Medieval Philosophy*, ed. N.Kretzmann (Dordrecht)
- O'Connor, D.J. 1967: *John Locke* (New York)
- Passmore, J.A. 1986: 'Locke and the Ethics of Belief', in *Rationalism, Empiricism and Idealism*, ed. A.Kenny (Oxford)
- Peacocke, C.P. 1983: *Sense and Content* (Oxford)
- 1986: *Thoughts: An Essay on Content* (Oxford)
- Popkin, R.H. 1983: 'The Third Force in 17th Century Philosophy: Scepticism, Science and Biblical Prophecy', *Nouvelles de la République des Lettres* 1:35–64
- Putnam, H. 1975: *Philosophical Papers II: Mind, Language and Reality* (Cambridge)
- Quine, W.V.O. 1963: 'Two Dogmas of Empiricism', in *From a Logical Point of View* (New York)
- 1965: *Word and Object* (Cambridge, Mass.)
- Ramée, P. de la 1572: *Dialectica* (Paris)
- Reid, T. 1785: *Essays on the Intellectual Powers of Man* (Edinburgh) [Scolar Press facsimile]
- Rogers, G.A.J. 1978: 'Locke's *Essay* and Newton's *Principia*', *Journal of The History of Ideas* 39, 2:217–32
- Rorty, R. 1980: *Philosophy and the Mirror of Nature* (Oxford)
- Russell, B. 1956: *The Analysis of Mind* (London)
- 1959: *The Problems of Philosophy* (London)
- Ryle, G. 1958: *The Concept of Mind* (London)
- 1968: 'John Locke on the Human Understanding', in *Locke and Berkeley*, ed. C.B.Martin and D.M.Armstrong (New York)
- Sacks, O. 1989: *The Man who Mistook his Wife for a Hat* (London)

- Schankula, H.A.S. 1980: 'Locke, Descartes and the Science of Nature', *Journal of the History of Ideas* 41, 3:459–77
- Schiffer, S. 1987: *The Remnants of Meaning* (Cambridge, Mass.)
- Smiglecki, M. 1638: *Logica* (Oxford)
- Spencer, T. 1628: *The Art of Logic* (London) [Scolar Press facsimile]
- Stalnacker, R.C. 1984: *Inquiry* (Cambridge, Mass.)
- Stewart, M.A. 1979: 'Locke's Mental Atomism and the Classification of Ideas (I)', *Locke Newsletter* 10:53–82
- 1980: 'Locke's Mental Atomism and the Classification of Ideas (II)', *Locke Newsletter* 11:25–62
- Strawson, G. 1989: *The Secret Connexion: Causation, Realism and David Hume* (Oxford)
- Strawson, P.F. 1964: *Individuals* (London)
- Suarez, F. 1866: *Disputationes Metaphysicae* (Paris)
- Wisdom, J. 1957: *Philosophy and Psychoanalysis* (Oxford)
- Wittgenstein, L. 1958: *Philosophical Investigations*, ed. G.E.M.Anscombe and R.Rhees, trans. G.E.M.Anscombe (Oxford)
- 1964: *The Blue and Brown Books*, ed. R.Rhees (New York)
- Wright, J. 1983: *Hume's Sceptical Realism* (Manchester)
- Yolton, J.W. 1956: *John Locke and the Way of Ideas* (Oxford)
- 1970: *Locke and the Compass of Human Understanding* (Cambridge)
- 1984: *Perceptual Acquaintance from Descartes to Reid* (Oxford)

Index

- Aaron, Richard I. 278n26, 290n36,
291n61, 292n65
- abstraction 42, 49f, 85, 213, 266f;
see also universals
- accidents, real 25, 55, 67, 272n50, 273n53;
see also substances
- affirmation 16, 71, 90, 109, 180, 266,
272n33;
see also judgement
- after-images 253
- algebra 80
- analogy 98, 100, 263
- analysis *see* method
- analytic and synthetic propositions 87,
194f, 225;
see also knowledge (trifling and
instructive)
- angels 76f, 106
- animals' thought 39, 124, 127, 155f
- Anscombe, G.E.M. 177, 245f, 287n49
- appearances, always true 29f, 145, 286n25;
see also ideas (simple and complex),
signs
- a priori* thought 62, 66, 120, 122, 125,
194f, 232, 239, 243, 244, 248, 255, 266;
see also abstraction, universals
- Aristotle, Aristotelian philosophy 16f, 22f,
23, 31, 31, 38, 52, 55, 59, 70f, 76, 86,
196f, 198, 213, 237f
- arithmetic *see* number
- Arnauld, Antoine 18, 33, 45, 46, 51, 59f,
74, 76f, 203, 213;
see also Port Royal *Logic*
- Ashcraft, Richard 106f, 283n155
- aspect-shift 154f, 163;
see also 'seeing as'
- assent 17f, 70f, 90;
see also judgement, belief, probability
- atheism 105, 279n64
- atomism 11, 14f;
see also mechanism
- Augustine, St 23, 29
- Bacon, Francis 14, 77, 98
- Bayle, Pierre 203, 207
- behaviourism 14, 63, 170f, 177f, 211, 244,
250, 256, 266, 288n50;
see also functionalism
- being idea of 42, 138, 221f;
scale of 45;
see also existence
- belief, opinion 70, 90, 109, 127, 135f,
237f, 240, 277n13, 281n85;
and desire 170f, 175f;
see also assent
- Berkeley, George 43, 57, 158, 161, 165,
201, 207, 213, 220f, 225, 230f, 252
- blindsight 123, 160, 172, 174f
- Bold, Samuel 140
- Boyle, Robert 11, 14f, 26, 102
- Buridan, Jean 179f
- Burnet, Thomas 294n104
- Burthogge, Richard 294n104
- Cambridge Platonists 196, 204, 207;
see also More
- categories 16, 19f, 59f, 74f, 273n53
- Caterus 45f, 50
- Catholics and Protestants 105f, 136f
- causality, perception of 135, 139f, 144,
158

- cause and effect, idea of 141;
see also power
- children's thought 127, 156, 221, 234
- Chomsky, Noam 86, 256, 260, 263, 280n72
- Chrysippus 285n1
- Churchland, Paul M. 286n30, 287n47
- Cicero 29f
- classification 102, 218f, 228f, 237;
see also abstraction
- clocks 101, 104, 203;
see also mechanism
- coexistence 67, 84, 87f, 140f, 158, 166,
 280n78;
see also necessary connection
- colour 22f, 25f, 33f, 50, 160, 177, 183,
 191, 228, 288n63, 295n116;
see also primary and secondary
 qualities
- common law 261f
- common sense 22;
see also ideas (of two senses), sensory
 field
- communication 216, 237;
see also meaning, understanding
- compound terms 31f
- concepts 13f, 151, 153, 286n30, 289n20;
see also ideas, pretheoretical thought
- consciousness 13f, 62, 83f, 123f, 133, 170,
 177, 181, 211f, 244, 266, 287n42 *et
 passim*;
see also intentionality, evidence,
 knowledge (primary and secondary)
- copula 16, 19, 73, 179f;
see also particles, affirmation
- Culverwel, Nicholas 293n99
- Davidson, Donald 281n85, 287n49
- defeasibility 110, 133, 147, 152, 267
- definition 86, 117, 242;
 real and nominal 31f, 86;
see also ideas (clear and distinct)
- demonstration 43, 70, 71, 76, 80, 83f, 89,
 216, 222, 263, 279fn64, 292n73f,
 295n116
- Dennett, Daniel C. 287n46
- Descartes, René 12, 16, 22, 27f, 29, 32, 33,
 38, 40f, 43f, 45, 47, 51f, 56f, 59f, 71,
 72f, 76, 77, 78, 80f, 83, 98f, 104, 106,
 110, 127, 135, 138, 145, 151, 158, 165,
 170, 196, 196f, 198, 203, 204, 206, 213,
 221, 233, 239, 255
- diagrams 43f, 81, 213, 222, 292n73f;
see also demonstration
- Digby, Kenelm 24f, 31, 60, 67
- Diogenes Laertius 29f
- distinctions of reason 25f, 213f, 221,
 272fn50;
see also abstraction, substances (and
 accidents)
- division of linguistic labour 237f, 240
- Dummett, Michael 276n142
- enthusiasm 40, 95, 105f, 107, 125f, 266,
 282n117, 283n151;
see also revelation
- Epicurus, Epicureans 12, 14, 29f, 32, 59f,
 70, 72, 138, 145f, 214, 266, 286
- error 40, 72, 75, 82, 91, 127, 145, 146,
 223, 226;
see also ideas (association of), will and
 belief
- Escher 226, 258, 268
- eternal truths 220, 291fn63;
see also universals, abstraction
- ethics 11, 80
- Euclid 76
- evidence 12, 29, 72f, 80, 83, 110f, 122,
 135, 147, 149f, 152, 168f, 214, 217,
 232, 258, 263, 264, 266;
see also knowledge (primary and
 secondary, sensitive), intuition
- existence idea of, *see* being;
 knowledge of 84, 89, 138, 222,
see also knowledge (sensitive);
 objective, *see* intentionality
- experience and theory *see* pretheoretical
 thought
- experiment 78, 99;
see also method
- faith 72, 74, 90, 106f, 280n84;
see also belief, revelation
- 'family resemblances' 119, 230f
- Field, H. 276n145

- Fodor, G. 276n145, 293n96
 formalism 263f;
 see also maxims
 forms 22f, 70f, 213;
 see also species
 foundationalism *see* knowledge
 (foundations of)
 freewill 127f;
 see also will, error
 functionalism 63, 133, 170, 244;
 see also behaviourism
- Gassendi, Pierre 12, 18, 28, 32, 38f, 41f,
 72f, 76, 135, 196, 199, 203, 205f, 214,
 221, 222, 263, 274n84, 274n93, 276n61,
 283n149
 Gettier, Edmund L. 283n163
 Gibson, J. 275, 277n13, 290n36
 Gilbert, William 99
 Glanvill, Joseph 38
 God idea of 45, 46,
 see also infinity;
 not a deceiver 28, 106f, 279n64
 grammar 255
 Grice, Paul 276n142, 288n62
- Hacking, Ian 270n7
 Hardin, C.L. 288n63
 Hartcliffe, John 294n104
 Hartlib, Samuel 106
 heat 145, 162, 167, 188;
 see also primary and secondary
 qualities
 hermeneutics 105f;
 see also revelation
 Hobbes, Thomas 16, 18, 23, 29, 38, 39, 41,
 60, 62, 73f, 76, 78, 80f, 83, 86f, 90f, 98,
 107, 135, 196, 197f, 200, 204f, 207,
 216, 222, 231, 237, 255, 267, 290n44
 Hume, David 45, 48, 102f, 129f, 159, 166,
 224, 267
 hypotheses 15, 77, 79, 98;
 see also method, analogy, probability
- 'idea', meaning of 12f, 17f, 38, 45, 52, 59
 ideas 12 *et passim*;
 adequate 33, 36;
 association of 57, 90, 97, 216, 223,
 225f;
 clear and distinct,
 obscure and confused 29, 34, 36, 40f,
 71, 206f, 239f,
 see also meaning;
 innate 27f, 57, 142, 144, 200, 257,
 see also knowledge (innate);
 as materials of knowledge 17f, 49, 52,
 53, 59, 135f;
 real and fantastical 32, 37, 106;
 simple and complex 14f, 16, 31, 41,
 106, 165, 238f;
 true and false 16, 30, 33, 59, 106, 165,
 183;
 of two senses 56, 165,
 see also sensory field
 identity 84, 209, 263;
 see also individuation
 illusion, misperception 65, 145, 152f, 163,
 167, 232, 251, 253, 289;
 see also error, 'seeing as'
 imagination 22, 38, 76, 96f, 130, 213, 216,
 220, 245f, 251;
 see also imagism
 imagism 14, 38, 59, 66, 212, 213f, 218,
 222f, 224, 245, 247, 252, 254 266f,
 295n116;
 see also intellect, imagination
 illumination 282;
 see also enthusiasm, evidence
 individuation 67f, 209, 229
 induction, principle of 98, 155
 infinity 41, 197f, 200, 203, 207
 intellect, pure 22, 38f, 40f, 59, 76, 213f,
 218, 221, 250, 254;
 see also imagism
 intelligence 232, 244, 255, 263f, 267, 296
 intentionality, intentional existence 22, 45,
 52, 59f, 62, 133, 150, 156, 170, 245,
 247f, 254, 266, 272n46
 intuition 76f, 89, 107, 125f, 240, 255f
 257f, 263, 267 *et passim*
 intuitionism 59, 83, 110f, 194, 228, 232,
 222, 224, 237, 243, 267f
- James, Susan 270n6, 276n3

- judgement 40, 16, 29, 71, 82, 90, 266, 277n13, 281n90;
see also affirmation, belief, assent
- justification 109, 111f, 114, 116, 118, 125
- Kant, Immanuel 28, 87, 96, 151, 203f, 224
- Katz, Jerrold J. 257, 259f, 262
- knowledge 70, 109 *et passim*;
 actual and habitual 74, 83, 122, 234f,
see also memory;
 classifications of 70, 80;
 definition of 49, 59, 90, 92, 138, 233f;
 degrees of 12, 74, 76;
 foundations of 35, 70, 145, 109, 110, 112, 121, 267f, 287n47,
see also defeasibility, maxims, pretheoretical thought;
 innate 121f, 232, 256, 263, 293fn99, 294n104,
see also ideas (innate);
 of one's own mental states 284n171,
see also reflection, consciousness;
 primary and secondary 121, 149f, 160, 172, 233, 248f, 267, 268,
see also reference, understanding (primary and secondary);
 relation to belief 70, 80, 83, 109, 135f, 266f;
 sensitive 12, 34, 53f, 80, 89, 133, 165f, 222,
see also perception;
 trifling and instructive 84, 86, 225,
see also analytic and synthetic propositions, definition (real and nominal);
 universal, *see* universals
- Kripke, Saul 186, 242
- Kulstad, Mark 274n99
- language reform of 80, 228, 230, 239, 241;
 and thought 16f, 53, 62, 127f, 155f, 176, 194, 216, 222, 237, 244, 251, 254, 255, 267,
see also nominalism
- Law of Nature 11
- Leibniz, Gottfried Wilhelm 34, 161, 201, 203, 207, 234
- Lennon, Thomas 283n151
- Lievers, Menno 275n137
- light 188f;
see also primary and secondary qualities
- linguistic competence 238, 241f, 255, 264;
see also understanding, language
- Locke, John *passim*;
 development of his thought 11, 20f, 200, 204, 214, 222, 292n73f
 logical form 62;
see also affirmation, copula, language (and thought), substances (and accidents)
- Lowde, James 294n104
- Lucretius 29f, 138
- Malcolm, Norman 288n49
- Malebranche, Nicolas 40, 45, 46, 52f, 106, 160f, 200, 203
- maps, mental 211f;
see also reference
- materialism 23f, 38, 39f, 170
- mathematics 11, 43, 71, 80, 229f, 264;
see also number, demonstration
- maxims 12, 71, 79, 234, 263f, 276fn8, 278n42
- meaning 63f, 67, 237, 251, 254, 266, 276n142;
see also language, understanding
- measurement 201
- mechanism 11, 22, 71, 88, 100f, 103;
see also clocks, atomism
- Mede, Joseph 106
- memory 22, 23f, 39, 41, 44, 66, 76f, 83, 214, 279fn64;
 primary and secondary 122, 248
- Mendel, Gregor Johann 117
- Mersenne, Marin 12, 70, 81
- method 17, 28, 76, 86f, 98, 116f
- millennarians 106
- Milton, John 273
- miracles 98, 102
- modes 74f;
 simple 19, 200f, 228, 238f;
 mixed 293n94
- Molyneux, William 19, 56f, 60

- Montaigne, Michel de 70
 More, Henry 196, 199f, 204f
 motives and reasons 128, 177
- Nagel, Thomas 287n45, 287n48
 natural belief 71, 129f, 135, 147f, 235
 natural kinds 186, 230f, 242f
 necessary connection 84, 87f, 103;
 see also coexistence
 necessity perception of 224, 232, 263f,
 267;
 and certainty 277n8
 negation 180f;
 see also privation
 Newton, Isaac 11, 102, 204
 Nicole, Pierre 76f
 nominalism 216, 222f, 225f, 228f, 272;
 see also language (and thought),
 universals
 Nozick, Robert 116, 284n169
 Nuchelmans, Gabriel 270n28, 280n84,
 288n51
 number 40f, 201, 221f, 224, 229, 256, 263,
 264
- observational predicates 152;
 see also private language argument,
 sensation (content of), pretheoretical
 thought
 O'Connor, D.J. 270n7
 optics 56f, 60
- pain 156, 161f, 163, 170, 178f, 189, 252,
 287n42
 Parker, Samuel 294n104
 Parmenides 70
 particles (grammatical) 19, 179;
 see also copula
 particular reference 64, 167f, 209
 passions 95, 127
 Passmore, John 91, 95f, 129f, 281n96,
 282n112
 Peacocke, Christopher 116, 151, 153f,
 284n169
 perception 132 *et passim*;
 and belief 146, 155f, 232;
 and concepts 151, 165;
 immediate and mediate 12f, 48, 60, 81,
 136f, 145, 151, 161,
 see also knowledge (foundations of);
 of particulars 64, 167f, 209;
 primary and secondary 123, 149f, 172f,
 175, 249,
 see also knowledge (primary and
 secondary);
 and sensation 146, 160, 165
 perceptual knowledge 146;
 see also perception, knowledge
 (sensitive)
 philosophical explanation 117, 125
 place 196f, 198f, 205
 Plato 23, 70, 77, 128, 196, 213, 237
 pleasure and pain 49, 137;
 see also pain Plutarch 285n1
 Port Royal *Logic* 18, 31, 71f, 76f, 135f,
 213f, 221, 239
 power 24, 26, 31, 33, 36f, 85f, 141, 190f;
 see also primary and secondary
 qualities
 precision 218f, 228, 237, 239, 260, 291n57
 predicables 26f, 213, 273;
 see also substances and accidents
 predication *see* affirmation, copula, truth
 pretheoretical thought 151, 158, 188f, 267,
 286n30;
 see also knowledge (foundations of)
 primary and secondary qualities 24, 50f,
 53, 98, 101, 139, 160f, 183;
 see also colour, heat, smell, power
 private language argument 177f, 183
 privations 33, 35
 probability 71f, 75, 80f, 90, 94f, 98, 109,
 281n110
 proper names 209, 242;
 see also reference
 proprioception 133, 162
 Protestants *see* Catholics and Protestants
 Puritans 96, 106
 Putnam, Hilary 237f, 243, 246, 294n109
- Quakers 106
qualia 48, 55, 156, 159, 177, 187, 245,
 249
 'quality', meaning of 55, 135

- Quine, Willard Van Orman 152f, 194f, 287n47
- Ramée, Pierre de la 76, 78
- realization 250f;
see also recognition, understanding
- 'reason', meaning of 105
- recognition 167f, 248, 268;
see also memory
- reference, primary and secondary 167f, 210;
see also knowledge (primary and secondary), particular reference
- reflection 42, 179;
see also consciousness
- Reid, Thomas 48, 161f, 190
- relation, as object of knowledge 85;
see also analytic and synthetic propositions, knowledge (trifling and instructive)
- Relativity, Theory of 207
- revelation 104;
see also enthusiasm
- Rogers, G.A.J. 282n137
- Rorty, Richard 151, 286n30, 287n47
- Royal Society 99
- Russell, Bertrand, 45, 84, 155, 245, 247, 254
- Ryle, Gilbert 84, 245, 270n7
- Sacks, Oliver 164, 295n136
- scale of being 45
- scepticism 12f, 27, 43, 51, 70f, 102, 106f, 135, 142, 184, 244, 266, 267f
- Schankula, Henry 278n26
- Schiffer, Stephen 276n142
- Scholastics 16, 45, 59, 280n84;
see also Aristotelian philosophy
- science 70, 216, 266f, 279n64;
 and natural history 11, 79, 88, 98
- 'seeing as' 65, 146, 154f, 163, 166f, 209, 251f
- senses, authority of 81f, 133, 135, 145, 266f
- sensation content of 64, 151, 187, 267;
 modularity of 286n30;
see also intentionality, perception
- sensation-reports 171, 177, 183;
see also observational predicates
- sensory field, unity of 133, 162
- Sextus Empiricus 29, 30
- signs, signification 16, 29f, 32, 44, 52, 60, 85f, 133f, 136, 139, 141, 170, 183, 216, 244, 267, 270n20;
see also ideas, intentionality
- simple apprehension 16, 18, 20f, 31;
 of propositions 18, 90, 270, 280fn84f;
see also affirmation
- smell 158, 162f, 287n36
- solidity 158, 167
- sound 163f
- space 194, 196;
 perception of 139, 158f, 160, 210
- species, sensible 22, 47, 52f,
 and genus 22, 26f, 213, 237
- spirits, place of 196, 200, 209, 290
- Stalnacker, R.C. 276n144
- Stillingfleet, Edward 139, 140
- Stewart, M.A. 270n13
- Stoics 29, 59, 70f, 214, 266, 294fn113
- Strawson, Galen 285n177
- Suarez, Francisco 25
- substance in general, idea of 42, 140
- substances and accidents 16, 19, 24, 31f, 67, 139, 213;
 ideas of 19, 31f
- Sydenham 11
- sylogism 17, 72, 76, 216, 255, 263
- testimony 83, 93, 98, 102
- time 194, 196, 203, 207
- toleration 94, 105
- topics 76f
- transubstantiation 25, 136f
- truth 16, 62, 64, 73, 216, 223;
 of appearances *see* appearances;
see also affirmation, logical form
- type-token ambiguity 13, 219f
- understanding general 41, 83, 122, 124, 232, 247, 251, 255, 267f;
 linguistic 237, 251, 254, 255,
see also language;
 primary and secondary 248, 254

- uniformity of nature 143f, 203f
universals, universal knowledge 11f, 42f,
46f, 70, 103, 209, 213, 255, 267f;
see also a priori thought,
demonstration, intuition
- vacuum 196, 204, 209
ventriloquism 163
Virgil 76
- Watson, J.B. 245
whole and part 71f, 79, 81
will and belief 17f, 70f, 73, 91, 127;
see also assent, error
Wisdom, John 261f
Wittgenstein, Ludwig 119, 133, 177, 183,
230f, 244, 245, 254, 276n142, 288n49
Wright, John 284n177
- Yolton, John 48, 280n78, 293n99, 294n104
- Zeno 232

Volume II Ontology

Introduction

The method adopted in the present work is explained and defended at some length in the Introduction in Volume I, to which the interested reader is referred. First, the work comprises an attempt to recover the meaning and original significance of the main philosophical argument of a single text, *An Essay concerning Human Understanding*, by reading it in the context of some of the ideas which were current or available at the time it was written, and which, on external or internal evidence, Locke certainly, probably or at least possibly had in mind when writing it. Second, the theory, or complex of interconnected explanatory and polemical models, which emerges from this interpretive enterprise is assessed in comparison both with its antecedent and contemporary rivals, and with later theory and argument, in particular, theories current today. Out of these comparisons, which are not always to the advantage of the present, a systematic response is developed and defended. In the course of comparison an effort is made to maintain a due sense of the historical distance at which Locke's text lies from us. The philosophical value of the exercise, and the light it may be expected to cast on the nature of philosophy itself, depends to a large extent on the very difference between the models and theories compared, and more fundamentally on the effort to avoid anachronistic distortion.

Virtually every paragraph in the *Essay* was directed towards a common end, that of forging an account of human knowledge, of what it is to know something and of what it is we can know, which would serve as a weapon against dogmatism in religion, morals and the study of nature. Locke's chief tool was the principle that the materials of knowledge come from experience, and he drew on the ancient empiricist tradition which had recently been brought to vigorous life by Gassendi and Hobbes. In ethics and religion his prime target was any irrational appeal to revelation, inspiration or conscience, and his earliest epistemological argument, in the work now called *Essays on the Law of Nature*, was directed towards putting ethics on a rational foundation. As was remarked in Volume I of the present work, his conception of what this involved had changed by the time of the *Essay*, but not his purpose. He did not try to separate morality from religion—just the reverse. Nor did he dispense with revelation. Rather, he tried to identify a central core of religious truth which is open to rational knowledge and from which morality can be deduced. Revelation is unnecessary within this sphere, unless for the benefit of those without the leisure or capacity to use their reason, and in any case all revelation is subordinate to reason: 'it still belongs to *Reason*, to judge of the Truth of its being a Revelation, and of the signification of the Words, wherein it is delivered'. In natural philosophy, on the other hand, it was not irrationalism, but too confident a rationalism which needed its wings clipping. Here the main argument was more like Gassendi's. The senses give us knowledge of the existence of things, but not of their substance or essence. Our speculations about their hidden natures are restricted to the ideas we get in

experience, and the best available hypothesis, that of atoms and the void, is evidently inadequate and incomplete. Our time would be better spent in careful observation and experiment, equipping ourselves to classify and generalize according to probabilities.

Although this volume is sequel to another, its argument is as far as possible independent and freestanding. Volume I was an attempt to reconstruct and assess Locke's theory of knowledge from the side of the mind, its contents, its 'operations' and its cognitive achievements. The concepts dominating the discussion were those of an idea, of relations between ideas, of knowledge, of belief and probability, of 'sensitive' and 'intuitive' knowledge, and of abstraction. Volume II, while dealing with some of the same general issues, will do so more from the side of the objects of thought and inquiry, their logical and ontological ordering, their classification and their individuation. Some points made in Volume I will be thought to bear repetition, and there will be occasional references back to earlier argument, but the dominant concepts in what follows will be those of substance, essence, mode and identity: concepts which appeared in Volume I only in subordinate roles. The discussion will lead into the topic of Locke's ethics, and his conceptions of law and the moral agent.

In Volume I, reason was found to defend certain theses more or less analogous to features of Locke's epistemology which are widely, if not universally, rejected today. These included a notion of the authority of perceptual knowledge, the corollary that perceptual knowledge possesses a hard boundary, a firm distinction between the level of speculative theory and the level of experiential knowledge, a strong analogy between *a priori* and perceptual knowledge and, perhaps above all these, an insistence that one of the conditions of knowledge (or at least, as it was argued, of 'primary' or paradigmatic knowledge) is that it is perspicuous or 'evident', so that the source of knowledge is apparent to the knower. When we have primary knowledge, there is no mystery to us as to why we believe what we believe. Knowledge is possible without this perspicuity, as when we can supply the answer to a question without knowing how we came to have it, or in the rather different cases of 'blindsight' and subliminal perceptual knowledge. But then it is secondary knowledge, parasitic on the possibility of primary knowledge.

Like many other seventeenth-century philosophers, Locke was directly influenced by ancient opposition to Scepticism. Like Epicureans and Stoics, he wrongly assumed that in the perspicuity of primary knowledge we can identify what the Sceptics' arguments had always seemed to call for, an unmistakable criterion of truth. Error he assigned to a faculty of judgement, rather than to the faculties which supply evident knowledge. Present-day theories of knowledge have correctly dispensed with the notion of infallible cognitive faculties whose deliverances are marked by some absolutely dependable sign. In consequence such theories, at any rate within the 'analytic' tradition, consist for the most part in attempts to explain and define knowledge without mentioning anything like 'evidence'—indeed to define it in terms which do not presuppose consciousness at all. It is this standpoint, as much as anything else, which sets seventeenth-century and present-day epistemology apart. Yet it is arguable (and was argued in Volume I) that neglect of the 'perspicuity' of knowledge, a condition only intelligible in terms of consciousness, is hardly less crippling now than the obsession with an infallible criterion was in the seventeenth century. What is needed is an epistemology which distinguishes the two,

allowing authoritative perspicuousness without infallibility.

The tendency for consciousness to fall out of consideration in presentday philosophy of mind has, of course, other sources. One, considered at some length in Volume I, is all that is wrong with Locke's explanation of the 'intentionality' of thought (i.e. its being about something, its capacity to be true or false) in terms of its being composed of elements which are in a special representative relation to things in the world. These elements are ideas, ultimately simple ideas which represent or signify their normal causes in sensation. They constitute the signs in a natural language of conscious thought, and are related by the mind in mental propositions just as conventional signs or words are related in the corresponding verbal propositions. There are good enough reasons, as it was seen, to reject this assumption that the logical form of sentences derives from the precisely correspondent logical form of the states of consciousness which they express, as if thoughts were a succession of sensations and images related as subjects and predicates. Nevertheless, neither these reasons nor others currently advanced are as cogent against the general and traditional assumption that it is conscious thought and experience, not language or action ('behaviour') or functional state of the brain which is the primary bearer of intentionality. This is not the place to repeat or even to summarize the arguments advanced in Volume I for reviving the claims of consciousness in this respect, although obviously they are consonant with the notion of the perspicuity of knowledge. One consideration there touched on, however, is closely related to the themes of the present volume. It was argued in more than one connection that, although no doubt our pre-linguistic experience and apprehension of the world cannot reasonably, or even intelligibly, be said to be itself of subject-predicate form, it is reasonable to assume that the possibility of primitive predication, and so of logical form, is generated by pre-linguistic capacities manifested at the level of consciousness, i.e. our animal capacity to pick out naturally discrete material objects ('substances') and our capacity to be struck by resemblances between them. Neither capacity presupposes language, but both, it would seem, are necessary to the development of language. Perhaps, too, both are necessary to that minimal grasp of the world which is a condition of any knowledge whatsoever.

That may look like a modest claim, as I believe in a way it is, yet in the context of twentieth-century philosophy it is possibly even more radical in its implications than the claim made in Volume I for the essential perspicuity of primary knowledge. While the latter would entail that currently fashionable approaches to epistemology merely eviscerate the notion of knowledge, and that tired old criticisms of classical intuitionism miss an important truth, to suggest that there are naturally discrete and unitary objects waiting to be picked out by us as the primitive subjects of predication is to offer to revive a form of realism which has been out of favour for centuries. An introduction is not the place to rehearse the argument for doing so, an argument which is carried on throughout much of what follows. Yet it may be appropriate at least to indicate the significance of the issue by taking some stock of the opposition.

What is at stake, it could be said, is a traditional distinction between distinctions: the distinction between real distinctions and distinctions of reason. According to this doctrine, very roughly put, substantial objects, things falling into the category of substance, are *really* distinct and unitary individuals, while entities in other categories—

qualities, quantities, events, actions and the like—are only *nationally* distinct, abstract, ideal individuals, ‘beings of reason’ sliced out of reality by the mind in virtue of predication. Kant was perhaps the first philosopher to hold that *all* our distinctions among objects are imposed or ideal, none are given or real. He held that even space and time are mind-dependent forms of our sensibility, and that we necessarily interpret the data of sense in terms of a set of categories which constitute the form of our capacity for unitary experience. In demonstrating what these ineluctable categories are, he took himself to be elucidating a conceptual structure which is in us *a priori* rather than a structure due to external reality. Among them, of course, is the category of substance. Yet the comprehensiveness and rigidity of his system made it as vulnerable as any dogmatic metaphysics. Non-Euclidean geometry, the Theory of Relativity and quantum mechanics have each seemed to disprove a central Kantian contention as to how we must think of an objective world. It is consequently not surprising if neo-idealists see the system of categories or conceptual scheme we employ as in principle subject to change. It is equally unsurprising if this contingent, mutable structure should no longer be attributed to the mind, as if it were a feature of human psychology, but directly to language, the medium of theory.

Twentieth-century analytic philosophy began in England with a vigorous rejection of idealism by Russell and Moore in favour of empiricism, but various influences soon established a strong conceptualist presence in the tradition. Wittgenstein and W.V.O. Quine, for example, can both be regarded, in their different ways, as heterodox neo-Kantians. One important argument for linguistic conceptualism, advanced by Wittgenstein in reaction against his own earlier doctrine, runs roughly as follows. If either presented reality or experience consisted in given objects, then (as his *Tractatus Logico-Philosophicus* had proposed) we could build up a language by simply naming the objects presented and devising resources for expressing relations between named items. Yet simply to point and utter a noise is not enough to introduce any term into a language. The context of a structured language itself is required in order even to determine what kind of word it is, what grammatical role it will play, and so what category of object it can be supposed to name. Without that context it is indeterminate whether the act of ostensive definition has picked out, for example, a book, its colour, its surface, its shape, some paper or whatever else lies in that direction. Unless sentential structure is brought to ostension, the utterance is at best an unstructured signal which picks out no object at all. Grammar itself cannot be taught by ostension, but must be learnt in the context of social, language-using life. It is, as it were, the form of such a life, determined by the complex purposes, practices and techniques which such a life involves.¹

Quine does not share Wittgenstein’s view of the significance and origin of logical form, but none the less offers a broadly similar, almost equally famous interpretation of ‘ostensive definition’. He argues in effect that the indeterminacy of reference cannot be resolved by repeated ostension in a variety of circumstances, or by its being determined in general under just what circumstances the utterance is ostensively appropriate. If we discover that ‘gavagai’ (a word in an otherwise untranslated language) is ostensively appropriate when and only when a rabbit is present, it does not follow that ‘gavagai’ means the same as ‘rabbit’. Although both words share what Quine calls the same

'stimulus meaning', the alien language might have a different logical form from ours, such that 'gavagai' denotes, not any kind of physical objects such as rabbits are, but 'mere stages, brief temporal parts of rabbits', or perhaps 'that single though discontinuous portion of the spatio-temporal world that consists of rabbits' (the 'rabbit-fusion'), or 'the recurring universal rabbithood' or, presumably, other objects beyond the capacity of our language to discriminate. For we cannot point to any of these objects without pointing to the others. To know that 'gavagai' denotes *rabbits*, it would be necessary above all to know what counts as the *same* gavagai, and that we have the same gavagai again when and only when we have the same rabbit again. In order to know that, it would be necessary to identify elements in the language to which the word 'gavagai' belongs serving the role of 'our own various auxiliaries to objective reference: our articles and pronouns, our singular and plural, our copula, our identity predicate'. Quine sums up a part of his doctrine in the slogan 'No entity without identity', and the approach to identity embodied in his argument will be discussed at a more appropriate place. What is relevant to present purposes is that the distinction of objects is taken to be imposed through the medium of a language with a particular and contingent logical form.²

Both Wittgenstein and Quine were influenced by, and reacted against, Russell's logical empiricism. Russell assumed without qualms that we can name objects given in sense-experience, but such objects are themselves elements of experience ('sense-data' comparable to Hume's 'impressions') rather than elements of an independent reality, and the notion of substance is held in low esteem. Another logical empiricist, the author of *Language, Truth and Logic*, ascribed the doctrine of substance to the crude projection of arbitrary linguistic structure onto reality: to 'the primitive superstition that to every name a single real entity must correspond', together with 'an accident of linguistic usage'.³ Like Russell and the early Wittgenstein, A.J. Ayer assumed the possibility of an ideal language actually respondent to the form of reality (which for Ayer just is composed of 'sense-contents'). It is this possibility which the later Wittgenstein and Quine both rejected, but there the similarity between them ends. For Wittgenstein ordinary language, rooted as it is in our way of life, may be in principle mutable, like the life itself, but it is beyond criticism, or at least beyond the kind of abstract and oblique criticism which he took to be characteristic of philosophy. Quine, on the other hand, has wanted to put the programme of analysis and language-construction onto a new footing, employing Russell's logic as a primary tool. Crudely, he has advanced the view that any choice between rival conceptual schemes, since it cannot be determined by reference to naked and uninterpreted reality or experience, must be made on such pragmatic grounds as predictive and explanatory force, economy or logical perspicuity. Just for such reasons, Quine supposes, the human race has adopted a scheme in which substantial physical objects are basic; but it is open to scientists and logicians to pursue refinement and even radical improvement and reshaping of that scheme. The tendency and, it seems, the point of this argument is to assimilate philosophical ontology to scientific theory, the role of the analytical philosopher to the role of the theoretical physicist. Nevertheless Quine sees the language we have as a starting-point, progress from which must necessarily be gradual. The ship can only be modified while still at sea. Not all conceptualists favour the same model. Thomas Kuhn, for example, prefers to see conceptual change as occurring in

revolutionary shifts from one scheme to another.⁴

Quine's argument about ostension has a second level or stage connected with his conservatism, i.e. the famous claim that, ultimately, no linguistic considerations at all are capable of resolving the indeterminacy of ostensive reference. His idea, roughly speaking, is that to ascribe the same conceptual scheme to speakers of different languages involves the judgement that various elements of each language individually make the same contribution to the meaning of utterances as certain elements in the other, but that this judgement is at best the simplest hypothesis tailored to produce the simplest or easiest translation. Evidence of meaning consists in behaviour associated with whole sentences or rather, when the 'auxiliaries to objective reference' are in question, total patterns of sentence-utterance. The meaning ascribed to one repeated element in the totality will cohere with, and so be variable with, the meaning ascribed to others. The employment of native speakers' identity-judgements as evidence that 'gavagai' means *rabbit* rather than *rabbit-part* will merely assume that some other element in their language means *is the same as* rather than *belongs with*. Yet, according to Quine, there is no structured 'meaning' behind, or other than, the behaviour so interpreted. Consequently there is nothing to make our choice of interpretive scheme right or wrong apart from simplicity, and 'we are unlikely to find a very alien culture with a predilection for a very outlandish' set of objects of discourse only because we are looking for a simple dictionary of translation.

A related argument, if one seemingly less dependent on behaviourism, brings Donald Davidson to reject altogether the notion of 'rival' conceptual schemes and so 'the very idea of a conceptual scheme'. For talk of radically different or changing conceptual schemes faces a certain problem. If two such schemes are genuinely different, they will be incommensurable. How then could the barrier between the two languages be surmounted, and from what point of view could we be supposed to see on both sides of it? Davidson argues that we cannot make sense of the supposed problem of translation. The very conditions for my belief that I have to do with a being who thinks and makes meaningful utterances are the conditions for my achieving a translation of its utterances into my language. It is only in so far as I succeed in relating its utterances to the world as I interpret it that I can verify the hypothesis that it is speaking meaningfully. The possibility of radical conceptual divergence, Davidson concludes, can therefore be dismissed from consideration. Yet this argument would seem at best to prove, not that we all have the same conceptual scheme, but that, if there were conceptual schemes incommensurable with our own, we could never discover it. In effect the untranslatable thoughts of beings with alternative schemes (or perhaps just the differences in logical form between their languages and ours) have joined uninterpreted reality beyond the bounds of meaningful description. Nevertheless the argument confirms

Davidson in a conceptualism with a built-in respect for natural language and, at the same time, with room for a programme of analysis devoted to revealing the logical form of natural language most perspicuously. Crudely, he no more believes in absolute ontology than Quine, but sees no alternative to our proceeding as if it were possible.⁵ Davidson's position represents one kind of dissatisfaction with the conceptualism which sees the world as an 'amorphous lump' receiving form from whichever conceptual

scheme happens to be employed to slice it up into objects. P.F.Strawson has advanced a deeper, more specific line of criticism, very broadly based on Kant's 'transcendental' argument for inescapable categories.⁶ Strawson's thesis is that every possible conceptual scheme must incorporate an ontology of enduring substantial objects in space as a condition of its possessor's having experience of any objects at all, or of distinguishing between itself and other things. Despite the antecedents of his argument, he rejects Kant's extreme idealism which, by placing space, time and the categories in us, as mere modes of our apprehension of things, rather sharply divides the objects of experience from totally mysterious 'things in themselves'. The result of Strawson's argument seems to be a position ambiguous between conceptualism and realism. On the one hand, he may seem to be saying that, unless the world were in fact divided into discrete, enduring material objects, and unless we could grasp how it was so divided, experience of any objects would be impossible. On the other hand, much of his argument seems to be conducted from within a broadly conceptualist understanding of individuation. It does not seem that individual substances are for him 'given' objects: it is rather that, unless my experience were susceptible to a certain pattern of conceptualization, it could not appear to me to be experience of objects, or as my experience. Moreover, Strawson has often seemed to assume that language is the medium of conceptualization and individuation. The philosophical position developed in the present work owes a deep, if somewhat general and indeterminate debt to Strawson's writings. Yet an aim will be to present an alternative to conceptualist theories of individuation, and to demonstrate that the substance-accident distinction at the heart of traditional ontology expresses an insight into what reality itself gives to the form of experience and language: an insight which modern conceptualism has lost without trace, and which it may not after all have been a part of Strawson's purpose to recover.

However that may be, to achieve an understanding of Locke's motives and presuppositions it is necessary to take absolute ontology seriously, and to step right outside the framework of modern linguistic conceptualism. That, I will suggest, it is not really so difficult to do once we have rid ourselves of certain twentieth-century inhibitions. Concrete, discrete material things, among which are the traditionally paradigmatic 'substances', are after all plausible enough candidates for the role of 'given' objects whose distinctness from other things is independent of thought. Moreover, the suggestion that a structured language is prerequisite to our individuating such objects is a paradox shaken (as Locke might have put it) by every infant with a rattle or dog with a rat. It is not, at any rate, flying in the face of common sense to suppose that such discrete material things as rattles and rats are *in themselves* distinct from their surroundings, that they are discriminated at the level of sensation by the higher animals as a condition of the sophisticated control of the world which such animals possess (something beyond the reach of a jelly-fish, perhaps, or Locke's oyster),⁷ and that they fall under a category basic in our language *for just those reasons*. This redness, this softness, this roundness, on the other hand (not to speak yet of such things as dog-fights, walks and avalanches), are not naturally unified and bounded individuals in themselves or in reality, but are individual objects of thought only in consequence of 'conceptualization' and the use of language. Nothing but an intensive training in modern philosophy could make the

traditional contrast between real distinctions and distinctions of reason actually difficult to apprehend. The serious historian of philosophy needs at least to understand, first, how it could have been drawn as a matter of course and, second, how it has since come to be regarded (if it is regarded at all) as without rational foundation. If the contrast can not only be understood, but can also be sustained, that would explain how the subject-predicate form is generated, on the one hand by our primitive ability to distinguish certain natural objects, and on the other hand by our capacity to be struck by resemblances between those objects. In the latter case words or, more specifically, predicates supply a means by which we can draw boundaries where resemblance and difference, although real enough and sometimes striking enough, draw none. Very roughly, we are enabled to treat as absolute what is a matter of degree. By the same means we are able to construct such objects as universal or particular qualities, actions, events and the like, notional entities to be contrasted with substances, the real or given objects of experience.

It is central to the philosophical argument of the present work that, although none of the old pre-conceptualist or realist explanations of the opposition between substance and non-substance are satisfactory just as they stand, they employ a number of fundamental insights which point to a more adequate account of it. If that account, as it is given below, is even roughly correct, then all forms of modern conceptualism embody a sweeping philosophical mistake. With respect to this whole issue the *Essay* is a text of considerable, perhaps unparalleled, importance. Locke accepted, stated and defended a strong contrast between substances, which are the real and natural unities, and non-substances, which are in some sense abstractions or creatures of the mind. Yet at the same time his efforts to combine this ontological stance with less traditional theory led him into arguments which contained the seeds of two very different doctrines: phenomenalist or reductive empiricism, and general conceptualism. Not that Locke even began to formulate or envisage either of those incompatible positions. Nor were they, as it was once popular to suppose in the case of reductive empiricism, implied by his principles without his realizing it. Yet the positions and arguments which he did adopt supplied weapons for later philosophers of both tendencies. In Volume I it was seen how his notion of a simple idea and its signification gave a foothold for Berkeley's views on the distinction between primary and secondary qualities and the relationship between sight and touch. It can also be seen how his reduction of our sensory apprehension of substances to the perception of the coexistence of qualities should have prompted the Berkeleyian thesis that substances are nothing but collections of ideas, or the rather more sophisticated Kantian thesis that the concept or category of substance has a purely formal role exercised in ordering the data of sense. The same thoughts were promoted by some features of his account of classification set out below, although neither was intended nor implied by that account. As a last example, his very specific concern to reconcile a belief in immortality with his neutrality on the issue of materialism led him to write an argument about personal identity which has come to be read and employed as protoconceptualist scripture.

It is very largely such connections with later philosophy that have given rise to Locke's unjust reputation as an inconsistent and indecisive thinker, lacking the ability or courage to draw out the implications of his own principles. Yet his arguments need first to be

interpreted, and deserve first to be judged, against what went before rather than what came afterwards. In that light the *Essay* appears as an amazing attempt to deal comprehensively and systematically with the philosophical issues of the seventeenth century from a single, if sometimes complex and subtle standpoint. We can learn most from it philosophically if we keep our eyes on its unity of purpose even as we are assessing arguments which, chopped out of context, seem capable of being read as the expression of ideas more familiar or congenial to us. Centrally relevant to that purpose was a theory of substance which appeared subversive to some of his contemporaries but which was traditional in a profound respect. For, with whatever qualifications, Locke accepted that reality itself gives us the most fundamental objects of our thought, if not those which are most clearly and fully known. If something like that claim could be satisfactorily reformulated, the implications for the theory of logical form would not be negligible.

Part I
Substance and mode

1

Introduction to Part I

The topic of substance, essence and accident was often touched on in the first volume of the present work, in discussions of logical form, of method, of perception (including the distinction between primary and secondary qualities, and the perception of the 'coexistence' of qualities) and of innate ideas.¹ The notion of substance is almost as pervasive and just as important in the *Essay* as the notion of an idea, with which it interlocks. Its interpretation there has been no less controversial. The account of Locke's argument offered below is in conflict with what has been the received view, and perhaps is still a common one, that for Locke 'substance' was something to be opposed to all attributes as pure logical subject to all predicates. The latter understanding of substance would make it, of course, unknowable and indescribable in principle. For to have knowledge of something can only be to have knowledge of its attributes, and to describe it must be to ascribe predicates to it.

The alternative interpretation advanced here, round which much of the argument of the present volume will be developed, is that for Locke, as for his predecessors and contemporaries, substance was opposed, not to all attributes, but to 'accidents' in a technical sense, and above all to observable accidents, i.e., in the case of matter, to sensible qualities and to the powers to cause or undergo sensible change. To know the 'substance' which underlies a given set of coexisting observable attributes would therefore be to know the fundamental nature or essence which, in interacting with perceivers and surrounding things, is responsible for the object's various appearances, its sensible effects on other things, and its own sensible alterations. Substance and essence are thus not distinct from each other. 'Substance' is for Locke unknowable only because this underlying nature does not fall within the limited scope of our cognitive faculties. It is not unknowable in principle. Nevertheless it should be said very clearly even at this stage (if only because a brief exposition of Locke's theory given by the present author in the past seems often to have been misunderstood in this respect)² that there is according to Locke's common usage a distinction between 'substance' and 'real essence', although not what the Scholastics or Descartes would have called a 'real distinction'.

That distinction, which will be explained below, is just one aspect of a complex and systematic attempt on Locke's part, following in the path of Descartes, Gassendi, Boyle and others, to reinterpret and redeploy the terminology and framework of Aristotelian logical and scientific theory, adapting it both to the ontology of mechanism and to his own anti-dogmatic epistemology. Such an enterprise may seem a strange one. Given that the mechanists were so conscious of having new things to say, as well as of a need to remove the rubbish, not least the verbal rubbish, they took to lie 'in the way to Knowledge',³ why were they (and Locke not less than others) so ready to parade their

doctrines in the uniform of the enemy, the terminology of the Aristotelian categories, the predicables, of 'matter' and even sometimes 'form'? Some commentators have been inclined to see in such survivals an uncritical inheritance, a signal failure to shake off the conceptual shackles forged by the very philosophy under attack. It is, of course, true that Aristotelianism supplied, as it were, the ordinary language of philosophy, taught in schools and universities as a matter of course. Yet the *Essay* is one of many works which warned against assuming that familiar language has determinate meaning. Locke's adoption of Aristotelian terminology in the face of his own warning must be supposed a deliberate policy, pursued indeed with self-conscious qualifications and disclaimers.⁴ Yet how could he both ridicule the words 'substance' and 'accident' and continue to employ them himself?

In part his motive was rhetorical. Expressions introduced into philosophy for the fictions of Aristotelian theory are found reference among the realities which actually perform the roles assigned to those fictions. The technique is reductive, identifying all that the ridiculed terminology could possibly stand for, while rejecting what it has been supposed to stand for. Malebranche remarked that the ordinary philosophy constitutes a logic affording useful general terms in which to discuss the natural world, and is objectionable only because of its claim that this abstract framework supplies detailed and specific explanations of things. The Aristotelians take abstractions for realities. That, of course, was the force of Molière's famous joke about the drug which works through its dormitive power, and Locke's jokes about the doctrine of substance were of the same kind. Malebranche deplored Aristotelian intolerance of those who, from the point of view of the new philosophy, take the trouble to 'attach distinct and particular ideas to these terms in order to understand' things.⁵ In the same spirit, Locke remarked sarcastically that,

were the Latin words *Inhaerentia* and *Substantia*, put into the plain English ones that answer them, and were called *Sticking on*, and *Under-propping*, they would better discover to us the very great clearness there is in the Doctrine of *Substance and Accidents*, and show of what use they are in deciding of Questions in Philosophy.⁶

He meant natural philosophy, and the immediate question at issue was that of the nature of space.

Indignation at the empty pretensions of Aristotelian philosophy was, then, joined with a certain respect for its abstract structure, and for the general intuitions and insights which that structure embodied. The present Part will constitute an attempt to say clearly how much of the Aristotelian theory of substance was honoured in the *Essay* and how far it had been transformed. Some philosophical assessment will also be offered of the ideas discussed, but much of the argument necessary for a more adequate assessment must be postponed to Part III, on identity.

Substance, essence and accidents before Locke

An account of the theoretical antecedents of Locke's discussion of substance, of the doctrines and arguments which it is essential for the reader of that discussion to bear in mind, has already been given in part. But some recapitulation and some filling out may be useful.

Substances or 'beings' (*ousiai*) are, on the Aristotelian theory, the fundamental entities in the universe, the ultimate objects of natural science; and that is their role in later theories, including (with certain reservations) Locke's. Substance is also the first of the logical categories, its ontological primacy being manifested in the logical principle that other categories are predicable of substance, while substance can be predicated of nothing else. Only an individual horse can be the subject of the substantival predicate 'horse', whereas the subject of the predicate 'brown' is not an individual brown but, for example, a horse. If such an individual as *this brown* can indeed be identified and made the subject of predication (as in 'this brown is fading'), it exists parasitically on a substance, e.g. for just as long as this horse is brown. It is logically distinct from the substance, but exists 'in' it. So it is, in one way or another, with entities in all other categories than substance, whereas substances exist primarily and in their own right. Non-substantial things are modes of the being of substantial things.

Such dependent existence was at least sometimes seen to carry with it a certain sort of mind-dependence.⁷ That is not because non-substantial beings do not in any sense exist independently of our thought, but because to pick them out and distinguish them as discrete entities or individuals is to make a distinction of thought or reason. It is not to point to a distinction which is already and independently there in nature. Individual substances, however, are really and naturally distinct, not merely logically distinct. As we might put it, for entities in all other categories, their individuality is dependent on our individuating them, but the individuality of substances is prior to their individuation by us. The horse exists as a separate individual in its own right, but its particular brownness is an individual which can be treated as a separate existent and subject of discourse only because it is sliced out of reality by an act of the mind.

An important discussion of substance occurs in the seventh book of Aristotle's *Metaphysics*, a book dominated by the question 'What is ousia?'⁸ We can roughly take this to be the question, asked of things identified by sensory means, 'What is ultimately there, as the real and independent object or objects of our ordinary beliefs and experience?' And this question Aristotle would have taken to be the question, 'What must we ultimately give an account of in objective, explanatory science?' He was concerned to identify the starting-points of scientific theory, the elements of reality whose natures constitute the fundamental explanatory principles in the universe.

Aristotle approached the question by way of the logical doctrine of categories. Neither 'being healthy' nor 'walking' denote an independent natural object (*kath'auto pephukos*), but 'that which is healthy' or 'the walking (thing)' (*to badizon*) does so. That is because such expressions as the last present something in particular, 'the substance and the individual' as the subject or 'substrate' (*hypokeimenon*). The answer to the question 'What is this thing, or substrate?' (*ti estin?*: What is it?—as opposed to 'What are its qualities?', 'How large is it?', etc.) is given by a specific substantive such as 'man' or 'horse'. Aristotle then raised the question whether any more ultimate answer can be given. He contrasted the normal view, that sensible 'bodies' (*somata*, a term which he understood to include animals, plants and their parts, together with air, fire, water and earth and their parts and composites) are substances in his sense, with rival Platonic conceptions of the explanatory principles of reality. Plato, of course, had advanced the doctrine that these principles are the 'forms' or 'ideas', universals to which he attributed independent, transcendent and undervived existence, existence in the primary sense.

Aristotle himself adopted a modified version of what he took to be the ordinary view. One modification was that parts of things are not genuine substances. The nature and definition of a hand or of blood depends on its relationship to the whole of which it is a part. Unless we know about that relation to something else we cannot know 'what it is'. Thus the universe as it appeared to Aristotelian science is populated by a multitude of different kinds or species of things and stuffs. These irreducible species together with their genera form a hierarchy which bestows a certain unity on Nature: as it came to be known, the great chain of being. Moreover the chemical kinds or stuffs are dynamically related according to a theory of 'substantial change', change in which one substance becomes another. Yet essentially each substance has its own nature or specific form and must be observed, studied and understood separately. Broadly, Aristotelian physics presents the sort of picture that is natural if biology and chemistry are regarded as fundamental and irreducible sciences.

For Aristotle, to give a scientific account of something is to define its essence. Its 'essence' (for which there was no single word in Greek) is what is necessary for its being what it is, or what it is to be that kind of thing. In a way, any possible subject of discourse has an essence, but Aristotle gave arguments for the conclusion that only substances, i.e. species and individual members of species, have essences in the full sense. (One such argument will be considered below.) Essences are always general, and the essence is identical with the specific form: substance and essence are one and the same. This latter principle extends to the individual substance, for the specific form *constitutes* the individual. Members of the same species not so much *have* the same essence, it seems that they *are* the same essence. In *Metaphysics* VII Aristotle at any rate seems to have thought that 'man' fulfils the same function of naming the universal essence of man whether it appears as subject of a general statement or as predicate of a particular one. 'Jack is (a) man' would then be taken to state an identity, not (as we might expect) between Jack and, indefinitely, some (particular) man, but between Jack and the form of man. This peculiar view might appear to lead to the consequence that all men are one and the same. Aristotle, however, seems to have held that the different 'matter' in which the same essence is embodied serves to keep individuals apart. Jack is the form of man

embodied in *this* matter, Mary is the form of man embodied in *that* matter. Although 'matter' serves in this way to individuate or distinguish particular things at one and the same time, it would be a mistake to identify the individual with its matter since the existence and identity of Jack is tied up with this continued embodiment of the form of man, not with the existence of the matter in which the form is embodied. The identification of the individual with the form (if that indeed had been Aristotle's view) was made rather less problematic by later Aristotelians who took care to distinguish the particular forms which constitute Jack and Mary from the universal form of man. The latter, they held, is an idea in the mind of the Creator like the general idea of a plough antecedently in the mind of a craftsman who makes a particular plough. In this way they managed to combine features of Platonic, Aristotelian and conceptualist theories of universals. On any version of the Aristotelian theory, matter cannot exist in the world without form, nor form without matter (with the usual, but not universal exceptions of the intellectual soul and intelligible forms).⁹

The formal structure of Aristotelian science was supplied by the Doctrine of the Predicables, the standard form of which derived from Porphyry's rewriting of Aristotle's logic. The five predicables were commonly illustrated in seventeenth-century logics along the following lines:

genus : animal
 species : man
 difference: rationality
 properties: the powers of laughter and speech, the
 possession of hands
 accidents : (separable) the coldness of cold water
 (inseparable) the blackness of crows

Scientific definition of a *species* is by *genus* and *difference*, so that the definition of the essence of man is 'rational animal'. Rationality, the difference, is 'the principle thing in a man's nature' and the *properties* flow from the difference 'as a natural emanation'.¹⁰ *Accidents* are attributes which are not thus connected with the essence: the substance can in principle be without them even when it never is in fact, as crows are never without blackness. The judgement that the blackness of crows is an inseparable accident rather than, in the full sense, a property, is the judgement, not just that a non-black crow or a crow's turning white is conceivable, but that we could fully understand a crow's nature without being able to infer its colour. A white crow would not be deformed, like a crow hatched without a bill.

The genus itself can be defined by genus and difference, i.e. by division of a higher genus, so that the ultimate or last species lie at the tips or twigs of an orderly hierarchical tree. The scientist aims at getting the tree right, basing his definitions on the careful observation of functioning. A definition which accords with the natural hierarchy, properly dividing a genus itself derivable from a higher genus by proper division (the highest genera of all being the categories 'substance', 'quantity' and so forth), was called,

for reasons which will soon appear, 'simple' or 'real'. Figure 1 represents the supposed real definition of *man* (see page 22).

A fully worked-out science consisted for the Aristotelian in a set of real definitions from which the properties have been derived. (There is no science of accidents as such.) Euclidean geometry can supply a model for such a science. Definition in geometry is not based on observation, however, since geometry is concerned with continuous *quantity* abstracted from change rather than, as natural science, with *substance*. Another point of contrast is the relationship between difference (or essence) and properties in the two categories: geometrical necessity is very different from the supposed necessity of the connection between being rational and having hands, which was teleological.

The doctrine of predicables was a theory, not only about the form of scientific explanation, but about predication, or the logical relationship between certain subjects and certain predicates. The name of the species is a term predicated of the individual in virtue of what constitutes the individual, its whole intrinsic nature, answering the question 'What is it?' When the name of the species is defined, that unitary nature is explicated, so that the predicate of the proposition 'Man is a rational animal' is simply the explication of the subject. It was commonly, although not universally held that the predication of 'properties' also fails to go beyond the subject. Only the predication of 'accidents' really adds anything to a subject.¹¹ This analysis casts some light on the notion of a 'naked substance', which is not a pure subject stripped of all attributes or grounds of predication, including those which constitute or are necessarily connected to its very being, but a substance stripped of all accidents for the purpose of determining 'what it is', and of identifying its properties and essence.

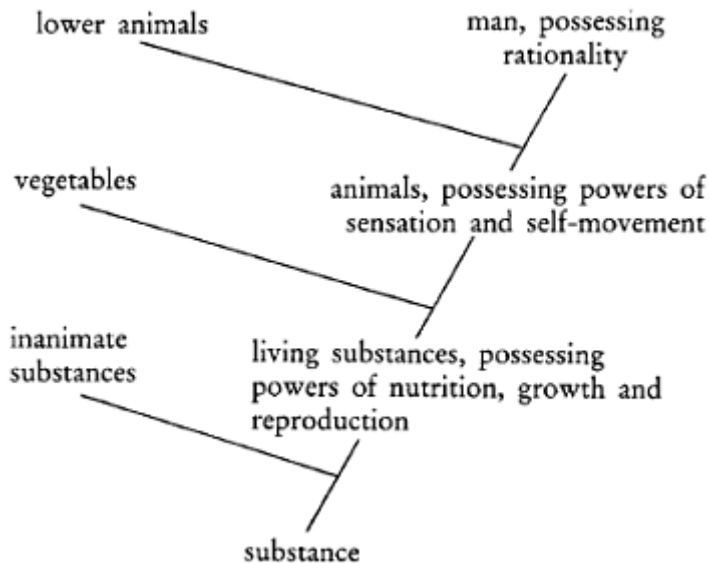


Figure 1 The tree of Porphyry

Scientific definition is opposed to two sorts of, as it were, pseudo-definition: the 'definition' of compounds, and the nominal 'definition' of simples. It is 'simple' by contrast with the former, and 'real' by contrast with the latter.

Compounds or 'syntheta' are named by such terms as are invented by combining items from different categories: in particular, substance and any type of accident. Their 'definition' thus involves composition rather than explication. Here are a few examples, some of which derive from Aristotle:

musician	– man who is musical
father	– male animal with offspring
palfrey	– horse which ambles
walker	– that which walks
<i>himation</i>	– fair man

Himation (actually Greek for a cloak) is a word coined by Aristotle for the purpose of illustration, and his reason for doing so was presumably to stress the infinite possibility of constructing such terms, the arbitrariness with which they can be invented and defined. There are no natural kinds corresponding to them, and so no genuine or ontological essences. Musicians (Aristotle's example) or palfreys (an example used in a hostile argument by Locke) do not present themselves as things of a kind, with a common nature to be named as such. A sign or corollary of this is that the individuality and being of the individuals which satisfy such predicates is in no way bound up with their doing so, i.e. if a man ceases to be musical, nothing substantial, not even a musician, ceases to exist. No individual is constituted by the 'essence' of a musician, since the musician *is* the man. Hence Aristotle's view that the 'essence' of a musician is not a true essence. But a man cannot cease to be a man and to have the essence of a man without ceasing to exist. This last argument raises issues which are among the most important in all metaphysics and philosophical logic, and its force and significance will later be explored in some detail in relation to Locke's own, very different views on species and individuals, on the classification and identity of substances.

The other type of pseudo-definition, nominal definition, may be of a natural kind, and may enable us to apply the name of a natural kind correctly by enabling us to identify the natural kind. Yet it is not scientific definition of the essence of the kind. For example, 'featherless biped with broad nails' is a nominal definition of the term *man*, since it is a description which fits man uniquely but which (it was generally agreed) fails to identify man's principal attribute or to effect a proper division of the genus, *animal*.¹² The definition of a compound term, of course, can never be anything but nominal, an explanation of the meaning of the word. To divide the class of men into those who are musical and those who are not, or water into that which is cold and that which is not, cannot be to reveal a natural essence or the natural articulation of the biological or chemical worlds. Once again, these are issues which were approached in a deliberately different way by Locke, but once again he defined his relationship to the tradition by adapting the traditional terminology to his ends.

One further element in the argument in *Metaphysics* VII is particularly relevant to the use made of Aristotelian doctrine by seventeenth-century proponents of the 'New Philosophy', and to the nature of their ambivalent reaction against the tradition. Early in that argument Aristotle suggested that, on the account of substance as that which is not predicated of a subject but of which all else is predicated, 'matter' might seem to be the ultimate substance. The passage very probably had a direct influence on seventeenth-century philosophy:

When all else is taken away evidently nothing but matter remains. For of the other elements some are affections, products and potencies of bodies, while length, breadth and depth are quantities, and not substances. For a quantity is not a substance; but the substance is that to which these belong primarily. But when length and breadth are taken away we see nothing left except that which is bounded by these, whatever it be; so that to those who consider the question thus matter alone must seem to be substance. By matter I mean that which in itself is neither a particular thing nor of a certain quantity not assigned to any other of the categories by which being is determined. For there is something of which each of these is predicated, so that its being is different from that of each of the predicates; for the predicates other than substance are predicated of substance, while substance is predicated of matter.¹³

Aristotle himself rejected this argument. 'Matter' is not a satisfactory response to the questions, asked for example about a man, 'what is it?' or 'what is that which is now walking?' 'Matter' identifies no particular thing, whereas the category of substance seems above all, as he put it, to pertain to what is truly separate and individual. Aristotle was, of course, aware of the theory of Leucippus and Democritus that the elements of the universe are immutable material atoms. Elsewhere he advanced logical arguments against the theory, but he evidently found its answer to the question 'what is substance?' quite implausible. It was contrary to his own scientific intuitions and, so he thought, to the common sense of mankind, which assumes that many sensible bodies, as such, are substances with irreducibly unitary natures.

It should be noticed in passing that for Aristotle matter does not even have in itself a determinate quantity. That is not, as it might seem, simply because he was strongly impressed by the logical argument that quantity is predicated of matter and is therefore distinct from it, but because on his physical theory quantity is tied to form. That is to say, matter is quantifiable only in respect of its form as, for example, a certain quantity of *air* or of *flesh*. In substantial change, when one substantial form is replaced by another, the matter of such and such a quantity of *water*, for example, may become the matter of such and such a different quantity of *air*, so that there can be no absolute answer to the question how much bare matter was involved in the change. Although '*materia prima*' underlies substantial change, it can be distinguished only in thought and could no more be quantified without the forms than it could exist without the forms.¹⁴

The great achievement of seventeenth-century philosophy was the assimilation of the truth which now seems axiomatic to most of us, that fundamental explanations in science

cannot start from the natures or essences of complicated physical objects such as are large enough to be perceived. No doubt we can and do make use of the notion of the 'nature' of a horse, but that nature, like the 'nature' of a machine, must be seen as dependent on its structure and the nature of the elemental parts which compose it. Only in that way can we think in terms of truly universal laws. Much can happen in a man or in a horse which is contrary to the complex and mutable, if in general fairly dependable nature of a man or a horse; but nothing can happen which is not 'natural' in the sense of conforming to the fundamental laws of nature. As Descartes put the mechanist analogy, 'a clock composed of wheels and counter-weights no less exactly observes the laws of nature when it is badly made, and does not show the time properly, than when it entirely satisfies the wishes of its maker'; that is, its end and nature as a clock.¹⁵

This profound change of emphasis was achieved in part by rejecting (or transforming) the Aristotelian conception of matter. On the Aristotelian view we can methodically identify accidents and strip them away in thought, leaving only the essence and the properties which flow from it, i.e. the naked substance. But to strip away the specific essence or substantial form is to be left with mere matter, something incomplete and natureless, incapable of distinct and independent existence. This the corpuscularians denied: matter has a nature and is therefore a substance in its own right. That is to say, matter cannot be posited even in thought without certain fundamental attributes, and in these attributes and the laws of motion lies the explanation of all that happens in the material world (except, perhaps, what is due to the will of spirits). For the Aristotelians, substances are the substrates of 'alteration', while pure matter is only the substrate, and a passive and ineffectual one, of 'substantial' change. For the corpuscularians (leaving spirit out of account), matter is the substrate of all change. There is no natural change except alteration, since substances do not naturally come into, or go out of, existence. The only intelligible alteration of matter is mechanical. The corpuscularians differed among themselves, however, on the question of which attributes are essential to matter and cannot be stripped off, and on the question of whether the laws of motion are necessary in themselves or arbitrary and contingent dictates of God. They also differed on whether we can know the answer to these questions.

For Descartes at any rate a secure answer to the first question can be given which is also delightfully simple: the fundamental attributes of matter which cannot be stripped away are extension, together with those properties (divisibility, mobility and so forth) which he supposed to be necessarily connected to extension. From the identification of the essence of matter with extension it follows that space is full and not an independent reality. In this respect Descartes' matter resembles Aristotle's, and in a certain sense it is also amorphous or undifferentiated in itself, although it is determined by motion rather than by substantial forms. Descartes was a 'corpuscularian', but his corpuscles are the minute parts of matter distinguished one from another not by empty space, the possibility of which he rejected, but by their vortical motion. The determinate shapes, structures and movements of any particular body at any particular time are one sort of 'accidents', since the body, i.e. that particular portion or quantity of matter, can survive their loss while retaining the essential determinable property, extension. Such determinate attributes Descartes called 'modes'.

The contrast between natural and merely logical individuals was taken over by Descartes, for whom there were officially three legitimate sorts of distinction. Firstly, a 'real distinction' can be drawn between substances, which can always be clearly and distinctly conceived as independent existents. Secondly, a 'modal distinction' can be drawn between, for example, a shape or a movement and the corporeal substance which possesses it; or between the shape and the movement of the same corporeal substance. That is to say, the substance can be conceived of as independent of a particular mode (such as its present shape or movement), and one of these modes of the substance can be conceived independently of the others, but no mode can be clearly and distinctly conceived of as independent of the substance to which it belongs. Thirdly, there are fundamental general attributes of a substance which we can distinguish in thought and from which we can distinguish the substance in thought, but which are such that neither they nor the substance can be clearly and distinctly conceived of in real separation from one another. Such 'distinctions of reason' can be drawn between the extension of a body and its divisibility, and between both these attributes and the body conceived of abstractly simply as a substance; yet the attributes are necessarily connected and such that we cannot form a clear and distinct conception of either without the other, nor of the body without both. The second and third distinctions, as Descartes himself made plain, can both be regarded as, in a broader sense, 'distinctions of reason' or merely logical distinctions, in contrast to the real distinction of substances.¹⁶

There is one determinate property, it should be said, which on Descartes' account a body cannot lose and which cannot be distinguished from the body's extension even in thought. Because Aristotle tied determinate quantity to form, he thought that to conceive of bare matter as ultimate substrate we must strip it of its quantity. For Descartes, to distinguish some matter, as such, from its quantity is (as the synonymy of the expressions 'some matter' and 'a quantity of matter' might suggest) an absurd error, for which he gave an insulting explanation. Apparent expansion must therefore be due to the addition of further matter, as a sponge appears to get larger when it absorbs water. The principle of its natural quantitative conservation (quantity being understood geometrically) is tied up with the very being and identity of matter. In effect, for Descartes matter just is substantial or real extensive quantity.¹⁷ Physics is the study of such quantity in motion, as Euclidian geometry is its study in abstraction from motion. The denial that matter can be stripped of its geometrical quantity is an important part of his argument that it has a nature of its own: a nature, moreover, of which we can have scientific knowledge. In this not every advocate of the 'New Philosophy' agreed with him: Kenelm Digby, for example, based his physics on the principle that matter can vary in density.

A second Cartesian category of accidents, which is ontologically reducible to the first, comprises sensible qualities as such. It is this sort of accident which, in the famous but confusing argument of the Second Meditation, Descartes imagined being stripped away experimentally from a piece of beeswax. Considering each of its sensible qualities in turn, smell, taste, colour, consistency and the like, he argued that the wax could survive the loss of that attribute when heated. Presumably because he was concerned only with its sensible or apparent features, Descartes included among its accidents even the quantity of the wax, since to observation heated wax expands. The immediate point of the argument

was that even to form an idea of the wax as something underlying observable change, as something barely extended and mutable, the intellect must interpret and go behind the deliverances of the senses. Hence imagists such as Gassendi and Hobbes are wrong to hold that the senses give us an idea of bodies, but no idea of anything incorporeal. For it is to the intellect, not to the senses, that we owe our ordinary idea even of a body such as wax.

It appears, however, that Descartes also intended the argument to be conducting us some way at least towards making explicit a fully clear and distinct, scientific conception of *matter*: towards stripping off its accidents and leaving it naked and exposed to our mental view. However that may be, the passage is a striking and important one both for the analogy which can be drawn with Aristotle's own discussion of matter as substrate, and for its firm association of the dichotomy between naked substance and the accidents of which it can be divested with the dichotomy between things as they are in themselves, and things as they are immediately presented to the senses. Indeed, in certain contexts, such as criticism of the notion of 'real accidents', seventeenth-century writers sometimes employed the term 'accident' virtually to mean *sensible quality*.¹⁸ Aristotle's assumption that some sensible things, as such, are substances with properties and essences open to observation had been roundly rejected. What we perceive by sense is nothing but accidents.

In accordance with the limited purpose of the Second Meditation Descartes elsewhere defined body as 'that substance which is the immediate subject of extension', while temporarily leaving open the question whether this formula effects a definition of the essence of body.¹⁹ The possibility is envisaged, that is to say, that 'extended substance' should be something like Aristotle's 'walking thing' or 'that which walks', where a more appropriate or fundamental answer can be given to the question 'what is it?' That possibility is closed for Descartes, however, when a fully clear and distinct idea of extension reveals it as indeed an independent nature (except in so far as it depends on God). Thus the distinction between extension and the substance which is extended is a mere distinction of reason, possible only because the formal concept *substance* applies to more than matter. This conclusion that there is no unknown subject of extension, i.e. that nothing further underlies extension, rests on the claim that Cartesian physics embodies a satisfactory and fully explanatory concept of matter. That claim was not simply an expression of satisfaction with certain mechanical laws, or of a prejudice (as it is often somewhat blankly proposed) in favour of what is quantitatively measurable. The reduction of all accidents to mechanical accidents offered an explanatory model of the substance-accident relationship: every accident is ultimately or in itself a determinate mode of a determinable property or essence, as roundness is a mode of extension. Any other conception leaves us with an unintelligible notion of 'real accidents' existing 'in' their substances in an unintelligible way.

The most important critical response to the argument of the Second Meditation was that of Gassendi, as set out in the Fifth Set of Objections. Gassendi agreed with 'what everyone commonly asserts, *viz.* that the concept of the wax or of its substance can be abstracted from the concepts of its accidents', but denied that this means 'that the substance or nature of the wax is itself distinctly conceived'. Admittedly, as well as 'the

colour, the shape, the fact that it can melt, etc. we conceive that there is something which is the subject of the accidents and changes we observe; but what this subject is, or what its nature is, we do not know'. He argued that what Descartes believed to be a purely intellectual conception of the substance of the wax cannot be free from the taint of sensory imagination. If we think of extension, we must think of it as limited, i.e. of its having determinate shape, and we can only conceive of shape as we perceive it. Moreover we have to call on sensible qualities, such as some sort of colour, in order to conceive of space-filling matter as distinct from bare space. The evident possibility of a real distinction between full and empty space, which Descartes denied, raised for Gassendi the question, 'what is that which is extended?': i.e. what underlies or explains the extension of a thing? ²⁰

Because of its relevance to Locke's arguments about substance, it is important to see that this last point has nothing to do with the notion of a pure logical subject capable of being stripped in thought of every last property, supposing the last of all to be bare extension. It is the sceptical demand for a more intelligible or explanatory account of matter's space-occupancy, in terms of what can be presented to sense and represented in imagination. That demand was to become a prominent theme of the *Essay* and in particular of the chapter on our ideas of substances. In general, Gassendi's claim that we can only conceive of the substance of wax in a confused way as an unknown 'something' did not derive from a muddled assumption that there are abstract subjects ontologically distinct from all predicates, but rested on his rejecting both the sharp distinction between sensory imagination and intellect and the pretensions of both Cartesian and Aristotelian physics. The most careful analysis of our experience of a piece of wax can achieve no more than an account of all the sensible qualities wax can take on, and changes it can undergo. This will supply the material for a definition of 'wax', 'but the alleged naked, or rather hidden, substance is something that we can neither conceive ourselves nor explain to others'. The result is a mere nominal definition, not a real definition. The process of stripping off the observable accidents of wax has not afforded us an informative conception of the substance of wax (or of body) distinguished from sensory irrelevancies. It has left us, on the one hand with a list of the defining phenomenal properties of wax, the characteristic powers and sensible qualities which lead us to call something 'wax', and on the other hand with a mere blank, corresponding to that of which we are ignorant.

Similar considerations apply to the soul as to body. Descartes found the essence of incorporeal substance in thought or consciousness, but Gassendi argued that the mind, like matter, can only be conceived of as an unknown 'something'. Consciousness is evidently the intermittent 'operation' of some substance, not the nature of 'the substance which performs this operation'. To mistake operation for essence suggests 'that you may be compared to a blind man who, on feeling heat and being told that it comes from the sun, thinks he has a clear and distinct idea of the sun in that, if anyone asks him what the sun is, he can reply: "It is a heating thing"'. ²¹ Gassendi surrounded his argument with what gave it force, namely difficulties in the notion of an unextended substance. It is the intelligibility or explanatory adequacy of the dualist ontology of Cartesian science which was at issue.

Gassendi's philosophy was nevertheless not blankly sceptical, at least as regards the

substance or nature of matter. The conceptions which we acquire in experience enable us to formulate a hypothesis, the theory of atoms and the void, which has at least some explanatory force and probability, however much it leaves unexplained and however impossible it is for us to prove it. This general view of man's epistemological relationship to the underlying principles of the natural world may exaggerate the dependence of theory on sense-experience, assuming that we have to think of the very small as if it was in principle perceptible. Yet it was perhaps in its time as important an intellectual advance as any aspect of Cartesianism. It at any rate supplied a theme which permeates the *Essay* and informed Locke's intricate and ambivalent attitude towards the claims of corpuscularianism. With the corpuscularians he scoffed at those 'who first ran into the Notion of *Accidents*, as a sort of real Beings, that needed something to inhere in, [and] were forced to find out the word *Substance* to support them'. But, like Gassendi, he was prepared to argue that not even the corpuscularians offered 'a clear distinct Idea of Substance', or gave a wholly satisfactory account of the substance-accident relation: thus 'of *Substance*, we have no *Idea* of what it is, but only a confused obscure one of what it does'.²²

Our complex ideas of substances and the idea of substance in general

The key to understanding Locke's general theory of substance is to realize that it is nothing other than a restatement and elaboration of the sceptical position adopted by Gassendi in his *Objections to Descartes' Meditations*, the position outlined at the end of the previous chapter. The elaboration is sometimes considerable, and to follow its twists and turns we need to refer to other antecedents than Gassendi and Descartes, but there is no shift in principle. Perhaps the best way to start on this thesis is with an examination of the opening sections of the chapter on our ideas of substances, a passage which has muddled critics and commentators from Bishop Stillingfleet and Leibniz to the present day. Here Locke described how we come to form ideas of substances, and arrive at a '*Notion of pure Substance in general*', on the basis of experience.

The chapter's title, 'Of our Complex *Ideas* of Substances', is itself significant. It begins with an explanation of that complexity:

1. The Mind being, as I have declared, furnished with a great number of the simple *Ideas*, conveyed in by the *Senses*, as they are found in exterior things, or by *Reflection* on its own Operations, takes notice also, that a certain number of these simple *Ideas* go constantly together; which being presumed to belong to one thing, and Words being suited to common apprehensions, and made use of for quick dispatch, are called so united in one subject, by one name; which by inadvertency we are apt afterward to talk of and consider as one simple *Idea*, which indeed is a complication of many *Ideas* together; Because, as I have said, not imagining how these simple *Ideas* can subsist by themselves, we accustom our selves, to suppose some *Substratum*, wherein they do subsist, and from which they do result, which therefore we call *Substance*.

2. So that if any one will examine himself concerning his *Notion of pure Substance in general*, he will find he has no other *Idea* of it at all, but only a Supposition of he knows not what support of such Qualities, which are capable of producing simple *Ideas* in us; which Qualities are commonly called Accidents. If any one should be asked, what is the subject wherein Colour or Weight inheres, he would have nothing to say, but the solid extended parts: And if he were demanded, what is it, that that Solidity and Extension inhere in, he would not be in a much better case, than the *Indian* before mentioned; who, saying that the World was supported by a great Elephant, was asked, what the Elephant rested on; to which his answer was, a great Tortoise: But being again

pressed to know what gave support to the 'broad-back'd Tortoise, replied, something, he knew not what. And thus here, as in all other cases, where we use Words without having clear and distinct *Ideas*, we talk like Children; who being questioned, what such a thing is, which they know not, readily give this satisfactory answer, That it is *something*; which in truth signifies no more, when so used, either by Children or Men, but that they know not what; and that the thing they pretend to know, and talk of, is what they have no distinct *Idea* of at all, and so are perfectly ignorant of it, and in the dark. ²³

The crucial problem of interpretation has been (and in the eyes of some commentators, still is) whether the notion of subject or *substratum* or 'pure substance in general' which Locke postulated as an ingredient of our complex ideas of substances was intended to be the notion of something unknowable in principle, i.e. unknowable for the *a priori* reason proposed by Leibniz in his critical response:

If you distinguish two things in a substance—the attributes or predicates and their common subject—it is no wonder that you cannot conceive anything special in this subject. That is inevitable, because you have already set aside all the attributes through which details could be conceived. ²⁴

On the other hand, is Locke's unknown subject unknown just because its nature is contingently unknown, a 'something' like Gassendi's 'substance or nature of the wax', or like his blind man's sun, a 'thing which heats' ?

Two questions are here worth asking. First, were there any other earlier or contemporary arguments, apart from Gassendi's, which were like Locke's argument and, perhaps, like it in ways in which Gassendi's was not? Second, does Locke's argument itself contain clues as to its point and purpose in the form of allusions, friendly or unfriendly, to identifiable philosophical doctrines? With respect to the second question, in the passage quoted we already have not only the technical terms 'substance', '*substratum*', 'accidents' and 'inhere' from the Aristotelian and post-Aristotelian tradition, but also the claim that our ideas of substances are complex rather than, as we are said to assume, simple. That 'we' include the Aristotelian philosophers whom Locke wished to controvert appears from his later reiteration of the charge in an explicitly technical, Scholastic way: 'These *Ideas* of Substances, though they are commonly called simple *Apprehensions*, and the names of them simple Terms; yet in effect are complex and compounded.' ²⁵ In l.xxiii.2, moreover, it is asserted that, as Gassendi had told Descartes, we do not have a 'clear and distinct' idea of the substance in which qualities or accidents inhere: not only 'Such Qualities, which are capable of producing simple *Ideas* in us' (i.e. sensible qualities), but the solidity and extension of 'the solid, extended parts'. There are plenty of indications here that Locke, whether or not struggling (as one commentator recently described him) to elucidate 'our conceptual apparatus', was certainly endeavouring to say things having a definite relation to Aristotelian theory, to Cartesian epistemology and to the corpuscularian theory of impenetrable particles. We do not fully understand Locke on substance unless we can explain the point of all this

polemical allusion, and can identify the place he wished to occupy in relation to the theories in question. Indeed, we can only get a firm grasp on what he was trying to say about 'our conceptual apparatus' through an explanation of the allusion, since the allusion is integral to his argument.

With respect to our first question, it is not difficult to find arguments which are like Locke's in having been directed specifically against the Aristotelians, in expressing a contrast between complexity and simplicity, and in being expressed in terms of 'substance' and 'accident'. Digby, for example, warned against the error of those who

confound the true and real natures of things, with the conceptions they frame of them in their own minds.... In the things, all that belongeth unto them is comprised under one entire Entity: but in us, there are framed as many severall distinct formall conceptions, as that thing sheweth it self unto us with different faces.²⁶

The same contrast, between the unity of the material object and multiplicity in our conception of it, or in its appearance to us, was drawn by Hobbes:

The Word *Body*, in the most general acceptation, signified! that which filleth, or occupyeth some certain room, or imagined place, and dependeth not on the imagination, but is a real part of that we call the *Universe*.... The same also, because Bodies are subject to change, that is to say, to variety of apparence to the sense of living creatures, is called *Substance*, that is to say *Subject*, to various accidents; as sometimes to be Moved, sometimes to stand Still; and to seem to our senses sometimes Hot, sometimes Cold, sometimes of one Colour, Smell, Taste or Sound, sometimes of another. And this diversity of Seeming, (produced by the diversity of the operation of bodies, on the organs of our sense) we attribute to alterations of the Bodies that operate, and call them *Accidents* of those Bodies. And according to this acceptation of the word, *Substance* and *Body*, signifie the same thing.²⁷

Finally, as we saw, the contrast exists in another claim of Locke's, made when he introduced his fundamental notion of a simple idea. Qualities that affect the senses are, 'in the things themselves', 'united and blended' but produce in us a variety of separate ideas which are 'simple and unmixed' with one another.²⁸ The thought is very like Digby's: 'But in my mind, every one of these notions is a distinct picture by itself.' It is not just noteworthy, but of the last importance for understanding Locke, that the original of this passage actually opens *Draft A* of 1671, and is immediately followed by the original of II.xxxiii.1.²⁹ Their juxtaposition provides about as direct and revealing an introduction to Locke's epistemology as he could possibly have offered. In *Draft B*, also of 1671, the equivalent passages (prefaced by an attack on innatism) are separated, but Locke still felt it necessary to append to his first explanation of our acquisition of ideas of sense and reflection a harbinger of his full account of substance:

Hence it comes to passe that we have noe Ideas nor notion of the essence of matter, but it lies wholly in the darke. Because when we talke of or thinke on those things which we call material substances as man horse stone the Idea we have of either of them is but the complication or collection of those particular simple ideas of sensible qualitys which we use to find united in the thing cald horse or stone...and which...because we cannot apprehend how they should subsist alone...we suppose they subsist and are united in some fit and common subject, which being as we suppose the support of those sensible qualitys we call substance or matter.³⁰

Note that the 'substance' or 'subject' of qualities is here matter by another name. No doubt this Hobbesian overtone disappeared from the *Essay* just because Locke did not share Hobbes' dogmatic conception of the essence of body: for Locke our idea of body itself includes the idea of an unknown nature together with the idea of those generic sensible qualities by which we distinguish body from spirit. As *Draft B* continues,

we have as cleare a notion of the essence of a spirit as any one hath of the essence of body. The one being supposd to be without knowing what it is, the substratum to those simple Ideas we have from without, and the other supposd, (with a like ignorance of what it is) to be the substratum to those actions or workings which we experiment in our selves within.³¹

Inevitably, then, Locke would have seen a need to distinguish the idea of matter from that idea of substance or *substratum* by which we think of the unknown essence or nature whether of matter or spirit.

There is another, related difference between Locke's argument and Digby's. Both were attacking Aristotelian theory, but different aspects of that theory. Digby was attacking the notion of 'real accidents' or 'real qualities', which he attributed to the mistaken, if natural assumption that the complexity or multiplicity in that everyday conception of a substantial thing which is founded on sense experience reflects a similar multiplicity in the thing itself: a multiplicity of accidents unintelligibly related to their substance. Digby did not doubt his own capacity to form a more sophisticated corpuscularian conception of what is out there, that would do greater justice to the real unity of the external object. Locke, on the other hand, saw the danger in a supposed natural inclination to *overlook* the complexity in our idea, an inclination stemming from our reasonable assumption that what is out there is unitary. What Locke was attacking was precisely the claim of Aristotelian theory that we can achieve simple notions of the essences of things. And indeed Aristotelian theory might well seem open to both charges. On the one hand, it was (or seemed to be) held that the variety in our sensory experience of an object simply matches an unreduced variety of sensible forms or accidents existing in a substance. On the other hand, the substance was supposed to have a functional unity which can be grasped by an act of *nous* or intellect working on the data of experience.

The hypothesis that Locke was intending to attack the Aristotelian doctrine of the

intellectual apprehension of an essence gains strong support if we compare II.xxiii.1 with a famous account of this process given by Aristotle himself. *Posterior Analytics* B19, a terse and corrupt but enormously influential passage, identifies (or is reasonably taken to identify) four stages in knowledge: (i) the perception of a thing or state of affairs, (ii) 'experience' or memory of repeated perceptions of similar things, (iii) the formation of a universal concept or thought, and (iv) the understanding which comes with a proper definition of an essence. In the case of substances (as it seems), (i) we perceive and (ii) we come to remember, as such, recurrent or repeated similarities between individuals; (iii) we arrive at a notion of a universal species (the same species on all these perceptual occasions) and finally, after more sustained observation and reflection has enabled us to pick out the genus and functional difference, (iv) we achieve a scientific definition of the essence of the species which explains the perceived concomitances of properties.³² There is a close structural correspondence between this process and the process Locke postulated which is unlikely to be coincidental. According to Locke, we do not form the complex idea of a substance as soon as we perceive, on one occasion, sensible qualities going together, but only after we have taken notice that they go 'constantly' together, i.e., as we should expect, after the second Aristotelian stage. The idea we then form (stage three) is universal and specific, but complex. 'Afterward' (stage four) we come to regard this idea as simple through a natural illusion. Locke was debunking the Aristotelian account of the process by offering a different interpretation of its last two stages.

Such a reading receives confirmation from the exchange between Locke and Stillingfleet about this passage, which Stillingfleet chose to understand as a monstrous claim that the general notion of substance arises, 'as a complication of many simple ideas together' from 'some fancies men have accustomed themselves to'.³³ He himself took the notion to be a general idea formed by intellectual abstraction. Locke's response was to deny that in II.xxiii.1 he had been concerned to attribute the general idea of substance to anything but abstraction, or indeed to make any point at all about that idea. He had been talking instead about our complex ideas of the species of substances, and these do involve a combination of simple ideas; yet because each specific idea includes the idea of 'a substratum, wherein that combination does subsist', the 'custom of supposing' such a unitary substratum misleads us into regarding the whole idea as simple.

This response did, however, leave one question unanswered: why exactly had Locke introduced the notion of 'custom' into his argument in the first place? He might have meant that habit so blunts our appreciation of what is involved in our specific ideas of substances that ('by inadvertency') we fail to realize that what distinguishes them from one another are the several ideas of their sensible qualities rather than ideas of their simple *substrata* or essences:

So that in this paragraph I only give an account of the ideas of distinct substances, such as oak, elephant, iron, etc. how, though they are made up of distinct complications of modes, yet they are looked on as one idea, called by one name, as making distinct sorts of substances.

Yet he went on to defend himself in a way which ascribes custom a role *before* the

supposition of a *substratum* rather than *after* it: 'I ground not the being, but the idea of substance, on our accustoming ourselves to suppose some substratum.' His defence is that this custom is indeed reasonable.³⁴ Why then did he write of 'custom' rather than of reason? Presumably just because the process he postulated was (like Aristotle's) quasi-inductive: we come to form the complex ideas which incorporate the notion of a *substratum* in response to the constant or repeated 'going together' of qualities and powers in our experience; just as Aristotle had held that the intellect forms the simple notion of a natural species on the basis of repeated remembered perceptions. Yet despite his self-explanation it is difficult not to see the reference to custom as doubly determined or dialectically unstable, looking both backwards to the constant conjunctions of stage two of the process and forwards to the 'inadvertency' of stage four. But however that may be, it is certain that a potentially mysterious feature of the paragraph, which Locke himself was prepared to represent as its chief point, actually consists in a polemical claim that a central Aristotelian epistemological doctrine is no more than a natural mistake.

In II.xxiii.2 Locke turned directly to the '*Notion of pure substance in general*' in order to emphasize that, in supposing a unitary substance behind sensible qualities, we have no idea of *what* we suppose beyond the idea of an unknown support of accidents. Where the Aristotelians had failed to achieve simple ideas of the essences of things no one else had succeeded. If the Aristotelian question 'what is it?' is asked with respect to the subject of secondary qualities, we can offer the corpuscularian hypothesis of solid extended parts or particles. But this is not an ultimately satisfactory answer, and the question arises again, what is the subject of extension and solidity? All we can say at this point is that it is 'something', using a word without a clear and distinct idea.

Locke thought that the question 'what is it?' is merely postponed by the best answer philosophers or scientists of the time could give, not because he thought that it would necessarily be postponed by any answer, but because he held (just as Gassendi had held with respect to Descartes' answer, 'extended substance') that the corpuscularian theory failed to capture the essence of what is there. Such scepticism is supported at length in II.xxiii in the argument that our idea of spirit is as clear as our idea of body. Since the argument is to the effect that our idea of body is as obscure as our idea of spirit, it can be read as ironical comment on the Second Meditation as well as a rebuttal of Hobbes' dogmatic materialism. We have 'distinct clear ideas' of certain properties of each, but not of the substance or fundamental nature of either.³⁵ In the case of body the postulated coherence of the particles remains unintelligible, which leaves their solidity and extension unexplained.³⁶ That is why the question, 'what is it, that that Solidity and Extension inhere in?' remains to be answered.

In II.xxiii.3 Locke moved on. First, he made the point that, although we employ the same 'obscure and relative *Idea* of Substance in general' in all cases, in forming our complex ideas of 'particular sorts of Substances' we suppose that in each case those sensible qualities involved in the idea flow from a 'particular internal Constitution, or unknown Essence of that Substance'. In other words, we see ourselves as postulating a *different* unknown nature for each distinguished sort or species, although of course we cannot specify what is peculiar to each of them except indirectly, in terms of the peculiar set of sensible qualities which enter into our definition of the sort in question: qualities

which, as Locke wrote with respect to iron or a diamond, ‘a Smith, or a Jeweller, commonly knows better than a Philosopher; who, whatever substantial forms he may talk of, has no other *Idea* of those Substances, than what is framed by a collection of those simple *Ideas* which are to be found in them’.

This point at least paves the way for Locke’s conception of a corpuscularian ‘real essence’ underlying the ‘nominal essence’ of any of our arbitrarily bounded kinds. As is explained at length in chapter 6 below, the ‘real essence’ of a species consists, on Locke’s view, in just those unknown modifications of matter which may be presumed to be responsible for the coexistence in nature of the qualities and powers by which we choose to define the species. Yet how much of this doctrine is already implicit in II.xxiii.3 is a matter of interpretation. The phrase ‘particular internal Constitution...of that Substance’ could be taken to refer directly to the way in which the postulated general substance, matter, is mechanically ‘modified’ or arranged. Nevertheless, it may be that a more accurate reading will take the argument here to be theoretically neutral as to the general character of the supposed specific essences.³⁷

However that may be, II.xxiii.4 and 5 seem clearly intended to saddle ‘us’ with conceptions of two general or universal substances: ‘*Body... the Substratum* to those simple *Ideas* we have from without’ and ‘*Spirit...the Substratum* to those Operations, which we experiment in our selves within.’ This movement in Locke’s exposition, especially if the notion of a general stuff subject to particular modifications is taken to be present in II.xxiii.3, would seem to have involved a certain shift in his employment of the expression ‘substance in general’. The argument of II.xxiii.5 is directed against the dogmatic materialist, and is to the effect that ‘from our not having any notion of the *Substance* of Spirit, we can no more conclude its non-Existence, than we can, for the same reason, deny the Existence of Body’. This argument is characterized at the beginning of II.xxiii.6 as having been concerned with ‘the secret and abstract Nature of *Substance* in general’, where ‘the nature of substance in general’ is presumably equivalent to ‘the nature of the general substance or substances’. Yet in II.xxiii.2 the phrase, ‘notion of pure substance in general’ had been employed for the general idea of substratum rather than for the idea of a general substratum. So a slide in the usage of ‘(idea of) substance in general’ seems to have occurred which may have been natural and innocuous enough in its context, but is potentially confusing to the reader.

Substance and real essence, matter and spirit, and the obscurity and confusion of the idea of substance

Three topics have been introduced which should now be pursued a little further: first, the relationship between the ‘substance’ of a kind of thing and its ‘internal constitution’, ‘real constitution’ or ‘real essence’; second, Locke’s somewhat ambiguous attitude towards mind–matter dualism; and third, the question of what precisely Locke meant by saying that our idea of substance is obscure and confused, i.e. not clear and distinct. These topics are, as we shall see, closely connected. The first two are related to important arguments elsewhere in the *Essay*, discussed in later chapters below. The third is related to Locke’s account of meaning, examined in Volume I of the present work. A fourth connected question is raised by a passage so far passed over, namely the appeal to language (i.e. to the linguistic or logical form of our definitions of substances) which appears in the second half of II.xxiii.3. Locke did not develop this line of argument until after the early drafts. Discussion of its genesis and significance will be postponed until the next chapter, below.

Substance and real essence

Sections II.xxiii.6–14 of the *Essay* constitute a single argument in which Locke explicitly brought together three ingredients of his theory: his characterization of the idea of substance as the unknown cause of the union of coexisting qualities and powers, his intimation that there is a general corporeal substance or stuff common to all sensible things, and his explanation of the secondary qualities and powers by which we chiefly distinguish the sorts of substances as due to the primary qualities of the minute parts of things. Towards the end of the passage there is something of a digression as Locke speculated on the reasons why God has not made the ‘internal constitutions’ of things open to our view, as perhaps they are open to the view of angels. The general purpose of the argument is clear enough, for it gives us all the ingredients necessary for the later distinction between real essence and nominal essence:

For our Senses failing us, in the discovery of the Bulk, Texture, and Figure of the minute parts of Bodies, on which their real Constitutions and Differences depend, we are fain to make use of their secondary Qualities, as the

characteristical Notes and Marks, whereby to frame *Ideas* of them in our Minds, and distinguish them one from another.³⁸

Nevertheless the argument raises a question about Locke's doctrine which it is not easy to answer, and is perhaps impossible to answer determinately. Briefly, the problem is this. In II.xxiii.6 it is stated that our ideas of '*particular distinct sorts of Substances*' consist in 'several Combinations of simple *Ideas*, coexisting in such, though unknown, Cause of their Union, as makes the whole subsist of itself'. It is this kind of idea which Locke called the 'nominal essence' of the species, the idea by which we distinguish, for example, iron from other things, represent iron to ourselves, and assign meaning to the word 'iron'.³⁹ Now the unknown cause of the union of the observable properties of iron, according to corpuscularian theory, is a particular or determinate constitution or modification of matter: precisely what Locke called the 'real essence' of the species. Yet in general Locke does not seem to have thought of the unknown substance and the unknown real essence of anything as identical. That is because he was understandably drawn towards saying that the unknown substance of, for example, gold and water, or a tree and a pebble, is the same, namely matter (whatever exactly that might in essence be). In other words, he regarded the unknown substance as an unknown general stuff.⁴⁰ Yet on his formal account of the idea of substance the reference of the idea ought to be the stuff-as-modified (if not the modification in question itself) rather than the stuff which is modified. Locke's usage uneasily spans both, in a way connected with the slipperiness of his expression 'idea of substance in general'. It is a tension which has left the relation between substance and real essence in his philosophy somewhat problematic, although not seriously so.

Some of the passages in which the two are explicitly distinguished have encouraged the unfortunate view that in Locke's mind, or in a part of his mind, the idea of substance is, after all, the notion of a naked logical subject distinct from all attributes, including essential attributes. Only in the 'real essence', it is supposed, did he offer something determinate enough to be knowable in principle, something which might figure in an ideal science and which is more than a grossly metaphysical 'x'. And it is true that he adverted more than once to what one modern commentator calls our 'twofold ignorance' of substance and of real essence.

At one place in the *Essay*, for example, he explained that our ideas of substances are inadequate on three counts. First, we cannot include in our definition of any species all the indefinitely many qualities and powers which every member of it will possess (i.e. the 'properties' of the species). Second, even that impossible achievement would not constitute knowledge of the real essence of the species. Locke's third point was briefly expressed: 'Besides, a Man has no *Idea* of Substance in general nor knows what Substance is in it self.'⁴¹ Another relevant passage occurs in the *Letter to the Bishop of Worcester*:

I do not take [the real essence] to flow from the substance in any created being, but to be in every thing that internal constitution or frame or modification of the substance, which God in his wisdom and good pleasure thinks fit to give to

every particular creature, when he gives a being...[And such essences] may be changed all as easily, by that hand that made them, as the internal frame of a watch. ⁴²

Several commentators have approved the view that in these two passages Locke conceived of substance (so clearly distinct from real essence) as a subject absolutely indeterminate in itself to which God might arbitrarily annex any attribute or essence he chose. ⁴³ Yet there is happily no need to ascribe to Locke such a metaphysically bizarre and uncharacteristic form of voluntarism. For the sentence quoted from the *Essay* may be roughly paraphrased: 'What is more, we do not even know the general or determinable nature or essence which constitutes all physical things—i.e. the fundamental nature of matter as such.' In the passage from the *Letter*, Locke was rejecting a suggestion by Stillingfleet that the 'internal frame and constitution' of each thing flows from its substance, i.e. from its specific nature or form (as the 'frame' of an axe was supposed to flow from its function). ⁴⁴ For Locke the real essence simply *is* the internal frame or constitution of the thing: more particularly it is those contingent modifications of the matter composing the thing which gives rise to the observable properties entitling it to be counted as of this or that humanly defined species. For such a real essence to be created, God must not only have created the matter or substance, but must also give it these or those determinate mechanical modifications. The latter are, of course, indefinitely mutable relatively to the substance. God could change them without altering the nature of the substance in any way. The matter would remain in essence matter whether it was so organized as to appear to us with the properties of lead, or with the properties of water: ⁴⁵ hence the analogy with a watch.

Locke's conception of the relationship between the general substance and its determinate modifications is perhaps even more clearly expressed in another passage from the Stillingfleet correspondence:

because the idea we have of the substance, wherein the properties of a man do inhere, is a very obscure idea: so in that part, our general idea of man is obscure and confused: as also, how that substance is differently modified in the different species of creatures, so as to have different properties and powers whereby they are distinguished, that also we have very obscure, or rather no distinct ideas of at all. ⁴⁶

Here it is evident that the distinction between the objects of our 'two-fold ignorance' is a distinction of reason, a Cartesian 'modal distinction', like the distinction between being extended and being square. Nothing could be less mysterious in principle than the relationship between the two: the mystery arises only because we are ignorant *both* of the general essence of matter *and* of its determinate modifications in this species or that.

So much for the relation between general substance and specific real essence. It should be noted that, although Locke distinguished the experiential properties by which we define a general substance, such as matter or spirit, from the unknown essence responsible for those properties, he did not mark that distinction by the terms 'nominal essence' and

'real essence'. These expressions are employed only with respect to the sorts or species of substances, and were reserved, as we shall see, for a very particular anti-Aristotelian purpose. In the case of the general substances, no distinction between substance and underlying essence could be drawn, and accordingly we find the terms 'substance', 'essence' and 'nature' employed interchangeably. What they refer to, however, could be opposed to the defining 'essence', or 'essence in relation to us'.⁴⁷

Materialism and dualism

Locke's argument in II.xxiii that we 'have as clear a Notion of the Substance of Spirit, as we have of Body' (having an obscure idea of both), is in a way sufficiently typical of its author, since it is an attack on dogmatic materialism from his moderately sceptical point of view. Yet it is also advanced from the point of view, or in the terms, of a dualism which, even if not dogmatically propounded, seems incompatible with principled scepticism. In this respect it is important to bear in mind two pieces of evidence in particular (there are many others) which underline the tentative character of Locke's substance-dualism: first, that other argument at IV.iii.6 to the effect that finite thinking things might, for all we know, be mere 'Systems of Matter fitly disposed'; and second, the gloss on both arguments which he supplied in the correspondence with Stillingfleet.

The main theme of IV.iii.6 may seem simple enough, but its allusions are complex and yet another problem of interpretation arises. Locke was saying that, because we do not know the essence either of spirit or of matter, we cannot know whether the human soul is material or immaterial. We know 'that we have in us something that thinks', but 'we must content ourselves in the Ignorance of what kind of Being it is'. Only if we knew the fundamental nature of things could we resolve such questions as whether a material mechanism could be conscious, or whether an unextended thing could exist. We do not understand how the former should be possible, but then we are equally in the dark as to how an *immaterial* thing could think, not to speak of its interaction with a material body.⁴⁸ Locke's point was that, although a dogmatic account of consciousness in material terms, such as Hobbes', was not (in the current state of knowledge) explanatory, his contemporaries had nothing better to offer and could not rule it out in principle that some such explanation might ultimately be given.

The Stillingfleet correspondence introduces further considerations. One is the status of animals, to which Locke's contemporaries (apart from Cartesians) were ready enough to ascribe consciousness, yet (apart from Cambridge Platonists) were slow to assign an immaterial soul.⁴⁹ Another part of Locke's case lies in the point that, whether finite thinking things are material or immaterial, they do not think essentially. As he argued against Descartes in the *Essay*, whatever it is that thinks in us, it is sometimes unconscious in dreamless sleep. Consciousness is to the soul 'what motion is to the Body, not its Essence, but one of its Operations'.⁵⁰ This model went with another: if God creates a substance, material or immaterial, he must also decide whether to put it into operation. Motion must have been 'superadded' to any moving thing, and thought to any thinking thing.⁵¹ We can now begin to see what is wrong with a popular interpretation of

IV.iii.6 (the interpretation apparently adopted by Leibniz) according to which Locke was saying that God might have superadded thought to matter by a kind of standing miracle, as Leibniz put it, overriding its nature.⁵² It is true that Locke agreed with the immaterialist that 'cogitation' is not to be found 'within the natural Powers of Matter'.⁵³ Yet all that that statement meant is that, like motion and other accidents, thought is not a 'property' flowing from the essence or nature of matter as such. Talk of quasi-miraculous superaddition would assume knowledge of what is and what is not unmiraculous (or 'natural' in a wider sense), whereas Locke's argument is purely sceptical: 'we know not wherein Thinking consists, nor to what sort of Substances the Almighty has been pleased to give that Power, which cannot be in any created Being, but merely by the good pleasure and Bounty of the Creator'.⁵⁴

A sceptical argument against dogmatic materialism may be perfectly concordant with a sceptical argument against dogmatic dualism. The question arises none the less whether Locke did not in fact allow too much to immaterialism in the argument of II.xxiii. Its conclusion, that it is 'as rational to affirm, there is no Body...as to say, there is no Spirit', seems to offer us dualism as the only rational position. Yet that implication was denied in the *Letter to the Bishop of Worcester*. All he had meant by 'Spirit', Locke there assured Stillingfleet, is *thinking substance*, 'without considering what other modifications it has, as whether it has the modification of solidity or no'. It is 'in the highest degree probable', but not demonstratively provable, that spirit in this sense is distinct from extended, solid, movable substance, i.e. matter. That both exist (i.e. both concepts are satisfied) is beyond doubt, but for all we know some things may satisfy both concepts.⁵⁵ Nevertheless this gloss seems to sit uneasily on parts of the original text. II.xxiii.5, after all, opens with the suggestion that the idea of spirit is formed precisely because 'we are apt' to presume a dichotomy of substance corresponding to the distinction between outer and inner sense. In xxiii.15, moreover, Locke explicitly repeated the argument with respect to the notion, not of *thinking substances* or *spiritual substances*, but of *immaterial substances*. What might seem particularly strange is that after the third edition he made changes to this section, actually strength-ening the impression that the idea of spirit with which he was concerned is precisely the idea of something which, by definition, is not material. The term 'spirit' twice became 'immaterial Spirit',⁵⁶ while several new sentences may seem to draw the argument even closer to Descartes' Second Meditation. The revised section ends:

For whilst I know, by seeing or hearing, *etc.* that there is some Corporeal Being without me, the Object of that sensation, I do more certainly know, that there is some Spiritual Being within me, that sees and hears. This I must be convinced cannot be the action of bare insensible matter; nor ever could be without an immaterial thinking Being.⁵⁷

There is, then, a certain paradox to be resolved. At roughly the same time as Locke was assuring Stillingfleet that his agnostic argument of II.xxiii.5 is free from the presupposition of metaphysical dualism (since it concerns the neutral idea of spirit as *that which thinks*), he seems to have been planning to strengthen the evidence for that

presupposition in the text itself. The explanation is perhaps something like this. In the first edition he did indeed attack dogmatic materialism from the point of view, and in the language, of the supposedly natural and probable thesis of dualism. This point of view provided a framework for the argument, but was not taken to be inescapable. Stillingfleet, however, was struck by the tension which exists between this way of arguing and the later attack on dogmatic dualism, and employed the seeming contradiction to embarrassing effect. Perhaps Locke's first reaction to this polemic was to purify his argument (or to claim that it was already pure) by moving to a consciously neutral interpretation of the concept of Spirit. That response made an important point: the dualist framework is unnecessary to the sceptical argument. Yet at the same time he did need to be discussing, not just *thinking substance*, but *immaterial substance*, since it was the conceptual possibility of the latter which was at issue with the dogmatic materialist. Moreover Locke's conception of God and our knowledge of God required that, even if we were (and knew ourselves to be) mere material machines, we should still have a use for the notion of immaterial substance in conceiving of God. That notion would still have to be acquired, if indirectly, through our experience of our own mental operations, regarded in abstraction from their (*ex hypothesi*) material basis.

Locke's proof of God's existence is an idiosyncratic combination of the cosmological and teleological proofs which will be examined in detail in chapter 14, below. All that is relevant here is that it is an argument to the effect that, from our intuitive knowledge of our own thinking, without regard to whether what thinks in us is material or immaterial, we are able to deduce the existence of an immaterial, essentially thinking substance as the cause of our thought. It is without much doubt this argument which influenced Locke's emendation of II.xxiii.15. The very curiously expressed thought at the end of that section as emended serves in effect as a qualification: even if I cannot after all be certain that the 'Spiritual Being' within me is not material, I can be certain that it could not have come to think 'without an immaterial thinking Being', i.e. without God. We know from the proof in IV.x that such code-phrases as 'bare insensible matter' expressed for Locke the notion of *matter unregulated by an intelligent being*. So Locke was not, after all, so inconsistent as to argue at one point in the *Essay* that mind-body dualism can be proved, and at another that materialism could be true. His position is expressed clearly enough in the Letter to Stillingfleet:

I grant I have not proved, nor upon my principles can it be proved...that there is an immaterial substance in us that thinks. Though I presume, from what I have said about the supposition of a system of matter thinking (which there [i.e. at IV.x.16] demonstrates that God is immaterial) will prove it in the highest degree probable, that the thinking substance in us is immaterial. ⁵⁸

Yet a problem remains, for why should Locke have thought that a demonstrative argument for the immateriality of God could also render it probable that what thinks in us is immaterial? Perhaps he believed that the traditional analogy between finite thinkers and the infinite thinker who made them in his image has some force. Even so, in view of the arguments which he was prepared to put on the other side, it is difficult to see how he

could have felt justified in speaking of the ‘highest degree’ of probability. It is tempting to suppose that he was here simply sugaring the pill of his agnosticism.⁵⁹

There were, however, other motives. His readiness to blow both hot and cold for dualism is at least loosely attached to his very similar attitude to Boyle’s corpuscularianism. In II.xxiii the two attitudes may seem inversely related, for the apology on behalf of dualism rests on a critique of the pretensions of contemporary physical theory. Yet that is a dialectical curiosity. For the attractions of Platonic dualism to Locke, as to many of his century, must have lain at least in part in the attractions of mechanism and of a mechanist explanation of secondary qualities. If colours exist only in the mind, it is attractive to suppose that there is a mind.⁶⁰ At the same time, as Hobbes’ system reminds us, the two did not stand or fall together. Even in Locke’s case the influence of corpuscularianism on his philosophy was far deeper than the influence of dualism, determining the very shape of his arguments about substance, real essence, species, identity and so forth. He was prepared to be scathing about Descartes’ theory of matter and sceptical about Boyle’s, but his distinction between a general substance and the determinate modifications of a substance was due to the former, while his employment of that distinction to explain Aristotelian species derived largely from the latter. Even the distinction just mentioned between the essence and the operation of a substance indicates how far the model of a machine dominated Locke’s thought: the potentialities of a clock lie in its structure, but it needs to be set going. More fundamentally still, his conception of physical explanation itself derived from mechanism. That is why his scepticism seems only to have extended to details of the theory. The world must be intelligible in something *like* the way Descartes had proposed, he assumed, something like Boyle’s theory must be true. Indeed Locke’s doubts often seem to have been concerned rather with the ultimacy or completeness of Boyle’s account than with its truth as far as it goes, or with its potentiality for at least partial explanations. It leaves certain things unexplained, such as the coherence of the particles themselves; but even in the heat of a sceptical argument Locke was apparently happy to assume that observable phenomena depend in the first instance on the mechanical interactions of minute particles. Mind-body dualism, on the other hand, was for him dispensable *in toto*. There may be dualist lapses, but he was in general prepared to frame his argument in carefully and provocatively neutral terms whenever (as in the proof of God’s existence or the treatment of personal identity) the issue had appeared to others to hinge on the question whether what thinks in us is material or immaterial.

The obscurity and confusion of our idea of substance

The third and last topic to be considered in the present section might be introduced by the following question. According to Locke, ‘substance’ means, roughly, ‘something which supports, or is the cause of the union of, coexisting sensible qualities’. This definition is intelligible enough, on present explanations. Why, then, did he hold that our idea of substance is obscure and confused, or even, as he sometimes suggested, hardly an idea at all? The idea may be relative, but he nowhere suggested that relative ideas are as such

obscure.

Here it may be helpful to examine four notorious sections of the chapter on space, II.xiii. 17–20, in which Locke can seem simply to have been berating the idea of substance in terms hardly compatible with his acceptance of its importance (even if *faute de mieux*) as a functional ingredient of an indispensable class of complex ideas. It has even been claimed that, whatever Locke was up to when he wrote them, these sections constitute a rogue passage which can safely be set aside while we seek an interpretation of more central passages.⁶¹ Yet it is unlikely that the argument here is anything but closely tied in with the argument of II.xxiii. Not only does II.xxiii.3 repeat the general point of II.xiii.18 that our idea of substance is not clear and distinct, but it does so with an explicit reference back to ‘the *Indian* before mentioned’ of II.xiii.19. Any interpretation of the one passage, it seems, must stand or fall with an interpretation of the other.

In II.xiii, faced with the question whether space is substance or accident, Locke (who seems finally, by the time of writing what became the first edition of the *Essay*, to have come round to a realist conception of space as *something*) responded that the question is unclear because we lack ‘a clear distinct *Idea of Substance*’. The argument is that to assert or to deny that space is a substance would be ‘to feign a Knowledge, where we have none, by making a noise with Sounds, without clear and distinct Significations’.⁶² Although the charge that we lack a clear and distinct idea of substance echoed Gassendi’s response to Descartes, the connection which Locke drew with linguistic confusion is better understood in relation to his own, quite unCartesian account of ‘clear and distinct ideas’ given at *Essay* II.xxix. Ideas are clear ‘when they are such as the Objects themselves...did or might, in a well-ordered Sensation or Perception, present them’,⁶³ while they are ‘distinct’ or ‘confused’ (at least, in the more interesting sense of these words) in relation to language. For example, if the terms ‘leopard’, ‘lynx’ and ‘panther’ are all indiscriminately associated with the idea (‘barely’) of a beast with spots, while yet we take them to be the names of different sorts of beast, then our ideas of leopard, lynx and panther are confused. That is, they are inadequate for making presumed distinctions.⁶⁴ It goes without saying why there is no ‘clear’ idea of substance for Locke, but the important claim in II.xiii is that the word ‘substance’ stands for no ‘determined’ or ‘distinct’ idea.

The argument is in effect as follows. God, finite spirits and body are all called substances. If we suppose that the word ‘substance’ as applied to God, finite spirits and body signifies a distinct idea, then these very different things must be supposed to be ultimately the same thing, i.e. they must be supposed, ‘agreeing in the same common nature of *Substance*’ to ‘differ not otherwise than in a bare different modification of that *Substance*; as a Tree and a Pebble, being in the same sense Body,...differ only in a bare modification of that common matter’. But that is ‘a very harsh Doctrine’ (in effect, like the theories of Hobbes and Spinoza). For we do presume a due distinction between the three which the idea of substance fails to mark. Locke then imagined the alternative proposal as being made: that the word ‘substance’ has three distinct meanings, standing for three distinct ideas. After sarcastically suggesting that in that case these three ideas had better be made known and distinguished by different names, he recurred to the claim that the term ‘substance’ lacks ‘clear distinct signification’. All this is entirely in line with

the view of II.xxiii that the function of the idea of substance is to stand for the unknown natures of things we discriminate from one another only by means of their experiential qualities and powers. In other words, the idea of substance in and by itself has no discriminatory force even where we can be sure that there is a real, underlying distinction (as between God and Body). *That is why it is 'obscure'.*

The next section, II.xiii. 19, is interesting in that it repeats the standard criticism of the Aristotelians, not expressed in II.xxiii, that they believed in a multiplicity of real accidents unintelligibly related to a substance distinct from them. The argument here does seem to be in tension with the account of II.xxiii in one respect. The word 'substance' is represented as if it were a mere technical term in a misguided doctrine, owing its existence to the naïve conception of 'real accidents'. Yet there is no serious incompatibility. The point is that we must recognize the Aristotelian model, for all its ontological pretensions, as an unwitting reflection of our state of ignorance:

we take it for a sufficient Answer, and good Doctrine, from our *European* Philosophers, That *Substance* without knowing what it is, is that which supports *Accidents*. So that of *Substance*, we have no *Idea* of what it is, but only a confused obscure one of what it does.⁶⁵

It was, as I have already suggested, a standard criticism of Aristotelianism that it merely formalized our ordinary, pre-scientific ways of talking and thinking and then took the result to be an explanatory theory. In something like this way, as we have seen, Locke held that the error of mistaking our complex ideas of substances for simple ones is both natural and Aristotelian. But the point extended beyond natural errors, and there is no reason to suppose that in II.xiii. 19–20 Locke was denying that our ideas of things and stuffs should properly contain the general idea of substance as a reflection of our ignorance of their fundamental nature. What he was denying is that the formal terms 'substance' and 'accident' are available to express anything useful about the nature of space. They are merely expressions by which the Aristotelians project the structure of our complex ideas onto things themselves. It is not that they are useless for all purposes, but that they are worthless 'in deciding of Questions in Philosophy', i.e. in natural philosophy.⁶⁶

All this was rejected in one hostile response to Locke's argument. For Leibniz, as for Descartes, the function of the general idea of substance is precisely not to discriminate among substances, so that its failure to do so is no mark of inadequacy. He saw the formal idea of substance as simply more abstract than the idea of this or that essence, not as the embodiment or manifestation of our ignorance of essences. For Leibniz, what substances have in common *qua* substances is not, of course, a common nature, but formal or categorial properties, such as the capacity for independent existence, or activity. From one point of view, Leibniz admitted, the idea of substance may seem thin and formal, but it will serve very well in a consideration of the ontological status of space (e.g. it could be said that substances, whatever their nature, are active, while pure space is necessarily inert).⁶⁷ Now whether or not Leibniz's certainly problematic account of the ontological status of space was preferable to Locke's scepticism, his account of the idea

of substance as a formal concept might well seem superior. Yet as will appear in the next chapter, Locke was no stranger to Cartesian views about the idea of substance, and had developed an interesting and subtle response to them.

Substance, mode and the argument from language

One of the most familiar of Locke's arguments on the topic of substance is the famous, or notorious, appeal to language. We should now be in a position to estimate its true significance. Locke wrote in II.xxiii.3 that

our complex *Ideas* of Substances, besides all these simple *Ideas* they are made up of, have always the confused *Idea* of *something* to which they belong, and in which they subsist: and therefore when we speak of any sort of Substance, we say it is a *thing* having such or such Qualities, as body is a *thing* that is extended, figured, and capable of Motion; a Spirit a *thing* capable of thinking; and so Hardness, Friability, and Power to draw Iron, we say, are Qualities to be found in a Loadstone. These, and the like fashions of speaking intimate, that the Substance is supposed always *something* besides the Extension, Figure, Solidity, Motion, Thinking, or other observable *Ideas*, though we know not what it is. ⁶⁸

What exactly was he claiming? At least, that we call gold, iron, horses, body, spirit, and so forth *things* (the term in ordinary language generally taken to be the equivalent of the philosophical term 'substances') just because we suppose them to have an unknown nature. Their names fall into the category of substance, and are noun-predicates rather than adjectives, just because our ideas corresponding to those names contain an 'obscure' element, the general idea of a *thing*. The noun-predicate, Locke thought, is unpacked by means of the suitably indefinite term 'thing' (in effect 'something') plus a relative clause. The same explanation is given of the traditional distinction between a substance, such as a loadstone or body in general, and the accidents said to exist 'in' the substance. Such a way of speaking intimates that we conceive of the loadstone or body as something more than those observable attributes by which we know it.

This epistemological explanation of a logical and grammatical category may at first seem merely odd and unsatisfactory, and we need to understand how Locke could advance it. Without such understanding, it is tempting to dismiss it as nothing more interesting than muddle. For however much we knew about a substance (even if we knew its most fundamental properties, its essence), it would still, it seems, be definable as a *thing* which possesses these properties. Something like this line of thought has contributed to the interpretation of substance-in-general as an entity unknowable in principle, a bare and entirely indeterminate subject of attributes. The appeal to language, it is thought, at least committed Locke to such a view. Yet that is to suppose Locke's

being forced to an absurd conclusion by a premise which, however natural it may appear to us, was explicitly rejected by him. For he held that the form of the definition of a substance term reflects the form of our inadequate idea of the substance. If we had an adequate idea, the definition would have needed to contain no place-marker. Hence he thought that we have a linguistic test for any pretension to know the essence of a substance. If the place-marker cannot be dispensed with (if the noun-predicate cannot be replaced by an adjective) the pretension is unfounded. For example, if the idea of extension contained all that is contained in the idea of body, then the two words would be inter-changeable:

we can never mistake in putting the Essence of any thing for the Thing it self. Let us then in Discourse, put *Extension* for *Body*.... He that should say, that one Extension, by impulse, moves another extension, would, by the bare Expression, sufficiently shew the absurdity of such a Notion. The *Essence* of any thing, in respect of us, is the whole complex *Idea*, comprehended and marked by that Name; and in Substances, besides the several distinct simple *Ideas* that make them up, the confused one of Substance, or of an unknown Support and Cause of their Union, is always a part: And therefore the Essence of Body is not bare Extension, but an extended solid thing; and so to say, an extended solid thing moves, or impels another, is all one, and as intelligible, as to say, *Body* moves, or impels. ⁶⁹

The 'essence in respect of us' is, of course, the complex defining 'essence', which reflects our ignorance. It is broadly equivalent to the 'nominal essence' of a species, but, as has been noted, that term was not available to Locke in the case of the general substance, body. The whole argument implies that the reason why 'extension' can stand neither as subject nor object of 'impels' is not because it is the nominalization of an adjective, but because it is the nominalization of the wrong adjective. If *x*-ness were what extension is not, the essence of body, then to say that one *x*-ness impelled another would make sense. The same argument, Locke continued, can be brought against the doctrine that reason is the essential property of man: 'no one will say, That Rationality is capable of Conversation, because it makes not the whole Essence, to which we give the Name Man'. ⁷⁰ In other words, Locke saw the ground of objection to the sentence, 'Rationality is capable of conversation', as not so much syntactic as semantic or, indeed, metaphysical. It is not intelligible that rationality should be the real essence of man, and that is why the sentence appears to us as nonsense. Rationality itself needs explanation: it is too evidently a mode or attribute dependent on some other attribute or attributes from which it flows. Similarly extension does not appear to us as an explanatory, self-intelligible nature. If it did, then its name could intelligibly replace the name 'body' by which, in our ignorance, we achieve indefinite and indirect reference to such a nature, the unknown *substratum* of extension.

Locke's appeal to language constitutes an explanation of the difference between noun-predicates and adjectives (or between substantival and adjectival nouns) which, however implausible it may at first seem to us, arose naturally enough out of the Gassendist

critique of Cartesian dogmatism. Indeed, it is against the background of explicit Cartesian commentary on the linguistic phenomena that his doctrine is most readily understood. Descartes, of course, had himself warned against distinguishing between a substance and its principal attribute:

Thought and extension can be regarded as constituting the natures of intelligent substance and corporeal substance; they must then be considered as nothing else but thinking substance itself and extended substance itself—that is, as mind and body.⁷¹

So the distinction between the notion of thought or of extension and the abstract or formal notion of substance itself is merely a conceptual distinction, and one which involves nothing but a greater degree of abstraction. In no way is the substance something more than, or other than, the principal attribute, and to think of it as if it were is to make a confused division of the substance from itself. Yet here Descartes' argument seems to have left a certain opening for anyone who rejected his dogmatic premise. For such an objector might have had the thought that our power of making just those distinctions which Descartes regarded as confused *without* falling into evident nonsense demonstrates that they are not, after all, examples of dividing a substance from itself. The point would seem all the more attractive in that those who adopted the Cartesian line sometimes employed the terms 'extension' and 'matter' interchangeably. For example, Malebranche advanced the argument that extension must be a being (i.e. a substance), since it is not a mode of anything in the way that roundness is a mode of extension.⁷² Locke noted this sort of usage and rejected it, together with the doctrine:

That *Body* and *Extension*, in common use, stand for two distinct *Ideas*, is plain to any one that will but reflect a little. For were their Signification precisely the same, it would be as proper, and as intelligible to say, *the Body of an Extension*, as *the Extension of a Body*; and yet there are those who find it necessary to confound their signification.⁷³

Malebranche's argument presupposes the Cartesian view that the metaphysical distinction lying behind the subject-predicate relationship is the modal distinction between a determinable essential attribute, such as extension, and determinations or modifications of that attribute, such as roundness. As he had put it in introducing the argument, the crucial question is 'whether matter does not have still other attributes, different from extension and those depending on extension', so that 'extension itself might not be essential to matter, and might presuppose something else that would be its subject and principle'.⁷⁴ After the explanation of the substance-accident relation in terms of Cartesian attribute and mode, however, he went on to discuss the distinction between extension and its subject as a merely verbal distinction without any clear semantic significance:

And what is said of [something else's] being the *subject* and *principle* of extension is said gratuitously and without a clear conception of what is being said, i.e. without there being any idea of it other than a general idea from logic, like principle and subject. As a result a new *subject* and a new *principle* of this subject of extension could in turn be imagined, and so on to infinity, because the mind represents general ideas of subject and principle to itself as it pleases.⁷⁵

The immediate context of this criticism was an attack on Aristotelian matter, whose proponents were alluded to as philosophers 'so accustomed to general ideas and entities of logic that their mind is more occupied with them than with those of physics'. The argument seems to be that, because Aristotelian matter is conceived of as *subject* to extension, a kind of essence or being is attributed to it which is distinct from extension. Malebranche complained that such an entity is not distinctly conceivable just because it lacks any attributes through which it might be conceived. Moreover, it is not only useless for explaining natural effects, it is also redundant, since extension alone is enough for physics.

What is interesting about Malebranche's argument for present purposes is what might seem least appropriate to a criticism of Aristotelian matter: i.e. the suggestion that the postulation of a subject to extension is the postulation of an *essence* to rival the Cartesian essence. This suggestion is inappropriate in so far as Aristotelian matter was standardly understood to be an abstraction, and would hardly have been said by its orthodox proponents to have an *essence*.⁷⁶ What we have, it seems, is a hostile characterization of the doctrine under criticism in terms its proponents would not have accepted, a characterization designed to present it as hopelessly confused. The result, however, might well have seemed a gift to someone like Locke, who *did* want to postulate an essence underlying the Cartesian one, an unknown essence. The point is this: Malebranche's claim that the concept of a subject of extension is a 'disordered abstraction', an abuse of 'the vague idea of being in general' seems to have committed him to regarding 'matter is extended' as a kind of solecism, as if the tautology ought rather to be expressed in the form of an identity, 'matter is extension'. Against this background, it is easy to understand Locke's insistence that 'Body is extended' is evidently *not* a solecism but the way we all have to talk, determined by the ideas we have. It is an attractive hypothesis that his reading of Malebranche in the early 1680s encouraged Locke to introduce elements into his argument not present in the drafts of 1671: not only the direct appeals to language, but the term 'idea of substance in general', later explained to Stillingfleet as 'the general idea of something, or being'.⁷⁷

Another book more than familiar to Locke which offered some discussion of the relation between substantives and adjectival nouns was the Port Royal *Logic*. This work explained predication and nominalization in terms of three kinds of objects of thought: things, things as modified (*choses modifiées*) and modes of things. Take as an example the proposition, 'this body is round'. The thought (or element of thought) corresponding to the subject is a conception of the body as a thing inhering in nothing else. The thought corresponding to the predicate is a conception of roundness as a mode dependent on the round body, which the *Logic* identified with a conception of the body as round, *une chose*

modifiée. The names of things, substantives, are 'absolute' terms which directly signify things in a distinct way. The names of things-as-modified, adjectives, have a dual role in predication. They directly signify or mark things in a confused way, and at the same time indirectly signify modes in a distinct way: Arnauld took it that 'round' in 'this body is round' is a name of this body conceived of as round, rather than, directly, a name of roundness, which it connotes or, as he put it, indirectly signifies. We can, however, employ a sort of terms which directly signify modes in a distinct way, i.e. adjectival nouns, or nominalized adjectives. These introduce modes as a direct subject of discourse, and because of the formal similarity of this function to that of primitive substantives, they are regarded as substantives from the point of view of grammar.

Now the argument took a turn more relevant to our present concerns. Most things are known almost entirely through the accidents or qualities which strike the senses. Hence they are known as *choses modifiées*, and our ideas of them are composite. They can therefore be divided: the idea of a mode or modes can be divided by nominalization from the idea of the thing. Presumably Arnauld meant that, for example, our idea of a ball is divisible into the idea of a body and the idea of a spherical shape, or roundness. Because most of our substance-ideas are divisible in this way (since we know and define sorts of substances through their sensible qualities) we become inclined to conceive even of simple things as divisible, 'dividing the essence of a thing into two parts, regarding one as the substance, the other as the mode'. Hence '*man* is often considered as the subject of humanity, *habens humanitatem*, and so as a modified thing'. This is a mistake, since it is to treat 'the essential attribute, which is the thing itself', as a mode, and as 'in a subject', a practice which leads to the employment of such adjectival, pseudo-modal nouns as 'humanity', 'corporeality' and 'reason'. The abstraction, it is implied, is illegitimate, and the scholastic nouns '*humanitas*', '*corporeitas*' and '*ratio*' are verbal monstrosities. The argument requires that, if Descartes was right about the essence of matter, the abstract noun 'extension' deserves to be placed in the same class, as a consequence of misguided abstraction or division.⁷⁸

Different as Locke's epistemological and metaphysical views were from those of Arnauld, it will be seen that some of them, at least, can be expressed neatly enough in Arnauld's terms. Locke held, not that just some, but that all things are known *comme choses modifiées*, so that the division of essences is always legitimate so far as *our* ideas are concerned. It would never occur to us, Locke was suggesting, to divide a substance from its true essence if ever we had knowledge of the essence. The terms 'extension' and 'reason' are for him, of course, entirely legitimate, and their legitimacy demonstrates that they are not true essences.⁷⁹ Nevertheless he recognized the illegitimacy of such scholastic abstract terms as are created by adjectivalizing the name of a substance and then nominalizing the result. But he had his own explanation of that illegitimacy, very different from Arnauld's and given in the short chapter 'Of Abstract and Concrete Terms'. The argument, as we should expect, is geared to his scepticism about essences, and it hinges on the bold hypothesis that the proper function of 'abstract terms' is to name known real essences. That is why ordinary language abounds with such words as 'whiteness', 'justice' and 'equality'. Conversely, our not ordinarily employing such terms as '*animalitas*' or '*humanitas*' constitutes

the confession of all Mankind, that they have no *Ideas* of the real

Essences of Substances.... And indeed, it was only the Doctrine of *substantial Forms*, and the confidence of mistaken Pretenders to a knowledge that they had not, which first coined, and then introduced *Animalitas*, and *Humanitas*, and the like.... Indeed *Humanitas* was a Word familiar amongst the *Romans*; but...was the abstract Name of a Mode, and its concrete *Humanus*, not *Homo*.⁸⁰

In this way Locke arrived at an explanation of the impropriety of abstract substantives which had epistemological implications quite the opposite of those of the Port Royal account. His explanation is, of course, entirely consonant with his view that the reason why the 'concrete' names of substances are nouns is because the ideas with which they are associated contain the general idea of an unknown nature. He dealt explicitly (and sensibly enough, if somewhat nonchalantly) with a potential counter-example: 'Some of the concrete Names of Relations, amongst Men chiefly, are Substantives; as *Paternitas*, *Pater*; whereof it were easy to render a Reason.'⁸¹ The doctrine that we know the real essences of simple qualities, modes and relations but are ignorant of the real essences of substances expresses the view that all *natural* essences are unknown. Other 'essences' exist as essences solely in relation to us. In the case of simple qualities this is because simple ideas each represent things according to one unanalysable way in which things affect the senses, and the import of such ideas does not go beyond that simple representative relation. In the quite different case of modes and relations, we can form complex ideas of them arbitrarily, without the need to refer to reality. We *know* their real essences just because we *make* them. The ideas of adultery, a triangle and so forth may be derived from nature but need not be; and in any case do not represent complexes of attributes supposed to enjoy a unity in nature. They are abstractions, even if sometimes abstractions from nature. Ideas of substances, however, are formed precisely in order to map natural things which exist as unities independently of our arbitrary conceptualization. The only way we have of doing such a thing is indirect and inadequate: we map such regular concomitances of various sensible qualities and powers as we take to manifest unified natural things. It is thus the different functions or purposes of our ideas of substances and our ideas of mixed modes which explain why a gulf opens up between the real and the nominal essence in the case of the former but not of the latter.⁸²

Some commentators have been puzzled by the doctrine that the nominal essence of a mode or relation is also its real essence. Would it not be better to say that a mode has no 'real essence', or perhaps that the terms 'real' and 'nominal' are equally inapplicable? As to the second suggestion, it has the disadvantage that it would remove one way of saying, as Locke wished to say, that our definitions of names of modes are entirely undictated by reality, and so solely a matter of the relation between ideas and language. Definitions of modes are always nominal. As to the first proposal, the meaning of the claim that nominal essences of modes fulfil the role of real essences has to do with the doctrine of predicables and Locke's philosophy of the *a priori* sciences. In effect, the distinction between ideas of substances and ideas of modes is Locke's explanation of the difference

between natural science or enquiry and the *a priori* sciences of mathematics and ethics. The latter are purely hypothetical, founded on arbitrarily constructed definitions. That is why they are possible *a priori*, without reference to natural facts. Yet they *are* sciences: 'instructive' and not just 'trifling' conclusions can be systematically derived, starting from their definitions, through the perception of relations. In the language of the predicables, 'properties' can be deduced from 'differences'. Given the definition or abstract idea of a *triangle*, which determines the denotation of the term, geometry can demonstrate indefinitely many truths ascribing further properties to triangles. That is so despite the fact that actual or supposed triangles need have no common nature beyond their triangularity, a superficial point of resemblance between them from the point of view of natural philosophy. Such a contrast between the science of natural substances and an abstract science of quantity was already present in the Aristotelian tradition, as an alternative to the Platonic view of mathematical knowledge as knowledge of transcendental universals. Given what Locke wanted to say, the doctrine that the nominal essences of modes and relations are their real essences was an entirely natural way of saying it.

Yet, if the real essences of substances are not their nominal essences, it seems that for sufficiently knowledgeable beings they might become so. Since access to real essences is in principle possible, it is possible that our knowledge of nature should take the form of a deductive science based on simple quantitative definitions, a universal mechanics:

to know the Properties of *Gold*, it would be no more necessary, that *Gold* should exist, and that we should make Experiments upon it, than it is necessary for the knowing the Properties of a Triangle, that a Triangle should exist in any Matter, the *Idea* in our Minds would serve for the one, as well as the other.⁸³

With such knowledge we could construct the ideas of determinate corpuscular structures (i.e. species of substances) without reference to their existence, deducing their behaviour in various possible circumstances. The analogy is with a locksmith who could tell by simple mechanics whether a key of such and such a shape would work a lock constructed in such and such a way. This does not, as it may seem, imply that all distinction between ideas of modes and ideas of substances, between abstract and natural sciences, would cease to exist. Locke's model allows that the latter would still be tied to reality as the former are not. It might be possible in his ideal physics to reason deductively to certain hypothetical conclusions about non-existent determinate structures, but for this to be *physics* the principles of our reasoning, the laws of mechanics, would have to be deduced from our knowledge of the general essence of matter, 'the Substance of Matter', a knowledge which could only be acquired by reference to the world. What Locke's appeal to language does imply is that in such an ideal state of knowledge we would not think or speak of that essential attribute as an *attribute*, or as the essence of something. That attribute, legitimately identified by an 'abstract', quasi-adjectival noun, would be what body ultimately *is*.

We might have expected Locke himself to have tried to approximate to such a way of talking in his most favourable (if tentative) commentary on Boyle's corpuscularianism,

when he seems to have been inclined to explain matter as, in effect, solidity variously spread about in space. In fact he employed the very expressions which on his own account ought to have withered away if solidity had been the known essence of body. Solidity is ‘inseparably *inherent in body*’, and is ‘the Idea *belongs to Body*, whereby we conceive it to fill space’.⁸⁴ According to Pierre Coste, Newton reported a joint speculation with Locke which seems to have been based on the identification of body with impenetrability. They imagined that we might get a rough idea of the creation of matter ‘if we suppose that God by his power had prevented the entrance of anything into a certain portion of pure space’, which would thereupon ‘be endowed with impenetrability, one of the essential qualities of matter: ...and we have only to suppose that God communicated this form of impenetrability to another portion of space in order to get a rough idea of the mobility of matter, another of its essential qualities’.⁸⁵ This speculation is interesting partly because it reveals something of the logical difficulties in identifying a thing with one or more of its attributes, however essential that attribute may be. For Newton was apparently forced to make *something* the subject of the predicate ‘impenetrable’, i.e. a portion of space. On the other hand portions of space do not move, so that he explained a body’s motion as the successive qualification of numerically different such subjects by the same attribute. One difficulty is to make sense of the possibility that ‘the same attribute’ should here mean ‘numerically the same particular attribute’. Locke himself saw substances as the primary place-occupiers, so that the identity of particular modes and relations over time is always for him dependent on the identity of substances.⁸⁶ If then, as Newton supposed, Locke was alluding to their joint speculation at IV.x.18, it is not surprising that it is there called a ‘dim and seeming conception’ of the beginning of matter, liable to put a strain on ‘Grammar it self’.

The difficulty of expression points towards a conclusion very different from Locke’s: namely that noun-predicates owe their difference from adjectives, not to our (in principle) remediable ignorance of natural essences, but to deeper, less contingent principles having to do with identity, continuity and existence. It is those principles, and their implications for logical form, which a satisfactory theory of substance must make explicit. It is not yet the place to reflect at any length on the metaphysical, logical and epistemological significance of the distinction between primitive noun-predicates and adjectival predicates. But it may be worth saying something now, in the light of the arguments currently under discussion, towards clarifying both that distinction and the philosophical issues which surround it.

First, however, some modern input deserves a mention. For the general suggestion that the syntactic category of a predicate can have deep philosophical significance has been widely disputed, and many regard it as a pure accident of language that ‘Jack is a dog’ cannot be idiomatically rephrased as ‘Jack is canine’, or even as ‘Jack canises’. There is no doubt that such expressions as the latter could intelligibly be introduced. The question is, however, whether there would be a loss in employing expressions which are formally adjectives or verbs in order to fulfil the function naturally and normally fulfilled by nouns; and if so, what kind of loss. What kind of stress, if any, would be set up between form and function? For in the end, of course, the considerations which lead us to distinguish predicates as nouns, adjectives or verbs are partly functional, if at the same

time partly formal or morphological. In the case of the distinction between primitive noun-predicates and adjectives the issue can be very narrowly identified just because there already exist such adjectives as 'canine' and 'human' distinct in meaning from the nouns from which they are derived. What is it that kept 'humanus' from meaning the same as 'homo'? Other examples illustrate the point even more compellingly: there is in English an adjective 'gold' (Latin *aureus*) as well as a noun-predicate of exactly the same form (*aurum*). The problem we are addressing could be put like this: why is it that 'gold' cannot always be construed as an adjective when it occurs as a predicate? What is the principle of distinction in this case?

Whether 'gold' is an adjective or a noun depends upon the category of that of which it is predicated. In 'this ring is gold' it is an adjective. In 'this quantity of stuff is gold', it is a noun. (In 'this sort of metal is gold', it is a noun too, but here the 'is' is the 'is' of identity.)⁸⁷ In the second proposition 'gold' is predicated of *the gold* here; in the first, of *the gold thing* here. *The gold* is the quantity of matter which may currently compose *the gold thing*, but which existed before it and will, unless annihilated, survive its destruction. Reciprocally, just as the adjective 'gold' is appropriately predicated of numerable objects ('things' in a somewhat narrower sense than Locke's, excluding 'stuffs'), so adjectives like 'canine' and 'human' may properly be used of quantities of stuff: a man is composed of some *human* flesh, bones etc., a plant of *plant* tissue. More generally, adjectives formed from the names of things or stuffs may be predicated of associated substances (including parts), but not of the substance named. Of course adjectives formed from both sorts of noun may be applied to a variety of subjects in categories other than the category of substance (the category of which 'things' and 'stuffs' are sub-categories). This may be done either in normal predication, as 'his manner was feline', or to form a noun-expression, like 'human geography'. In the phrase 'the gold stuff, however, the adjective 'gold' (assuming that it is not just a colour-word) is a barbarism, like 'canine' applied to dogs. Teeth may properly be called canine, but not dogs. Notoriously we tend to react to the artificial use of the adjective in the contrived expression 'human being' by treating the adjective as a noun. Etymological pedants who insist that we are 'human beings' rather than 'humans', as much as if they had insisted that 'gold' is always short for 'gold stuff', are working against a profound sympathy between grammar and meaning, form and function: they are not pedantic *enough*. Interestingly, Locke implicitly offered an explanation for this kind of case, for his account would suggest that 'gold stuff' is impermissible because reiterative; for him, in effect, the idea of gold already contains the general idea of stuff. One thing, at any rate, he was right about: the same logico-linguistic force both renders 'gold stuff' deviant and explains why the quasi-adjectival nouns 'humanity' and 'aureity', supposed to stand for what makes men men and gold gold, appear as barbarisms.

There are, of course, general nouns related to the primitive noun-predicates 'gold' and 'man' which can be said to name abstract objects: i.e. 'gold' and 'man' as they occur in such sentences as 'Gold is a metal' and 'Man is mortal'. But the abstract objects which such nouns name are the general types or species, not universal attributes or essences. In the case of thing-predicates, of course, the universal thought is more usually expressed by means of either the plural or the indefinite singular: e.g. 'Tigers are striped' or 'A tiger is

striped'. The status of substantial species and their names will be discussed below.

Adjectival or verbal nouns can play other roles than that of picking out universal attributes, and consideration of one of their functions should deepen our suspicion of abstract nouns like 'caninity', formed by nominalizing an adjective itself derived from a primitive noun-predicate. The expression 'this whiteness', or 'the whiteness of this thing', may be used to refer to a destructible particular in time and space. The same goes for verbal nouns. 'This blue is fading fast', 'The movement of the wheel has ceased', 'This walk is tiring', 'This walking is tiring', all fall into the same general group. Characteristically such particulars exist in so far as the predicate in question is satisfied by some *other* object. Yet even the difference between the nouns 'walking' and 'walk' indicates how rough such a characterization is. My recent *walking* and the contestants' *fighting* are particular activities, like stuffs in being quantifiable but not countable. *Walks* and *fight*s, like 'things', are countable stretches of those activities. That is not always just because they are continuous and discrete from other stretches of the same activity, but often because they are bounded by some other principle of unity, such as a deliberate intention or formal rules. Contestants may, deplorably, carry on with their *fighting* after the *fight* is over.

Entities like *battles* and *squares* and even *movements* may seem to achieve independence of the attribution of the corresponding predicates to subjects other than themselves. Yet that, to speak very roughly, is when the relevant predicates are relational. Objects squarely related, as well as square objects, may underlie the existence of particular squares. A particular movement can be conceived of as passing from one thing to another, but that is only when, on another and more basic conception of a movement, the particular movement of one thing causes the movement of the other. The things move in causally-related succession. The broadly Aristotelian principle holds: the existence-conditions of any such non-substantial 'object' are to be explained in terms of a general predicate's being satisfied by some *other* particular object or objects. What I here call 'existence-conditions' can be thought of as the spatial and temporal boundary-principles of such an object. Roughly, a battle lasts as long and extends as far as people are embattled.

Arnauld objected to the nouns '*corporeitas*', '*humanitas*' and even '*ratio*' on essentialist grounds: it is a blunder to separate a thing from its essence. A different, but related objection to 'caninity' and 'aureity' and the like can be put in the following way: whereas the predicate of 'John is happy' can be nominalized to allow us to talk directly of the particular, John's happiness, no such transformation of the predicates 'dog' or 'gold' (or their Latin equivalents) can pick out any *other* particular than the particular which satisfies the original predicate. These predicates do not, that is to say, connote an attribute which could in turn, as a particular, become the subject of discourse. To 'construct' such a particular, *this caninity*, *this aureity*, is to construct something with existence-conditions indistinguishable from those of *this dog* or *this gold*. Predicates like 'dog' and 'gold', as Aristotle put it, simply tell us *what the subject is*. Those present-day philosophers who claim that the logical primacy of substances is an accident of our conceptual scheme or syntax permit or approve such artificial words as 'caninity' or 'canises' as a means of exploring the possibility of rival, if perhaps less convenient

schemes. On a more traditional view, the one defended in the present work, no such alternative exists. To put it crudely, the category of substance contains the natural objects from which any conception or knowledge of the world must start, reference to which cannot be jettisoned from any adequate and explanatory account of the world. Locke's position was a kind of mixture of the traditional view together with his radical explanation of substance-attribute logic as a mark of our ignorance. The latter thesis may seem reductive, for what could be more contingent than ignorance? Yet in truth it no more questioned the primacy of substance than did the Cartesians' use of 'extension' for 'matter'.

There is doubtless something wrong with Locke's contribution to the theory of substance-attribute or substance-accident logic, but my chief concern has been to demonstrate that his argument arose naturally and plausibly enough within the context of Aristotelian and Cartesian logic as an ingenious way of advancing his kind of scepticism about the fundamental nature of things. We need not suppose, for example, that Leibniz's well-known criticisms raised any issues or made any claims of which Locke would not already have been aware from Cartesian sources.⁸⁸ Moreover, it would be quite misguided to think of Locke's doctrine as a passing aberration in the history of philosophy. One of the most popular logic textbooks of the eighteenth century, published many years after the *Essay*, followed his argument closely.⁸⁹ Perhaps the most significant and illustrious employment of his doctrine was by Kant, who was prepared to argue *both* (with Locke) that the possibility of dividing subject from attribute indicates that the latter does not constitute the ultimate nature of the former, *and* (with Malebranche and Leibniz) that we can always in principle continue to divide subject from attribute:

Pure reason demands that for every predicate of a thing we should look for its appropriate subject, and for this, which is necessarily in turn only a predicate, its subject and so on to infinity (or as far as we can reach). But it follows from this that nothing which we can reach ought to be taken as a final subject, and that the substantial itself could never be thought by our understanding, however deeply it penetrated, and even if the whole of nature were disclosed to it.⁹⁰

So Locke's idea of substance in general became transformed into something he never intended, through the adoption of a premise which he would not have accepted. It became in Kant's hands the pure concept of a logical subject which exists in us solely as the intimation of the unknowable 'thing in itself'; and, in the hands of later idealists, an intimation of the Absolute as ultimate subject of all predication. It is perhaps partly because of this tradition, not to speak of the twentieth-century empiricists' reaction against it, that the point of the seventeenth-century debate has proved so difficult to recover.

Species and their names in the corpuscularian world

Locke's critique of the Aristotelian conception of species in many respects follows in the wake of Boyle, who attacked the doctrine of elements and other principles of Aristotelian and alchemical science in such works as *The Sceptical Chemist* and *The Origin of Forms and Qualities*. Nevertheless Locke's argument is philosophically richer, not only because of its context within a more fully worked out epistemology and theory of meaning than was available to Boyle, but also because of its elegant and polemically effective redeployment of the 'predicates'. Yet, apart from its importance within general philosophy, it was, like Boyle's, a significant contribution to the contemporary debate on the theory of natural classification. It very probably helped to bring about a crucial change of mind by the leading British taxonomist of the time, John Ray,⁹¹ and it continued to have some effect on taxonomical thinking throughout the eighteenth century. Locke's aim was not simply to rebut the pretensions of Aristotelian science, but to raise the whole question of the place of species and genera in corpuscularian nature.

His first argument against the doctrine of specific forms, however, was quite unoriginal. It was simply, in effect, the doctrine of abstraction in its ancient role of a theory of universals. Essentially similar arguments can be found in philosophers as diverse as Descartes and Hobbes. As Descartes argued, all universals are simply modes of thought; or, as Locke put it, 'General and Universal, belong not to the real Existence of Things.' Distinctions between universals, such as the distinctions among the five predicables (at any rate as applied to particular sorts of substances), are also mind-dependent. The hierarchy of genera and species, ascending to the various categories of being, arises only because the mind ascends by abstraction from man and horse to animal, *vivens*, body, substance 'and at last to *Being, Thing*, and such universal terms which stand for any of our *Ideas* whatsoever'. Thus 'this whole *mystery of Genera and Species*' is 'nothing else but abstract *Ideas*, more or less comprehensive, with names annexed to them'. The 'Rule, that a Definition must consist of *Genus* and *Differentia*' is rejected. The method of division of the genus by the difference merely extracts an arbitrarily chosen element from the abstract idea of the species, one property out of all those which are contained in the complex nominal essence.⁹² A further point brought against the tree of Porphyry seems to be this: that hierarchy would have to be founded in the form or essence of the individual, since universals exist for the Aristotelian only in individuals. Consequently there must be a corresponding complexity or layering in the individual. Thus the Aristotelian must 'think Nature to be very liberal of these real Essences, making one for Body, another for Animal and another for a Horse, and all these Essences

liberally bestowed upon *Bucephalus*'. This is to mistake the 'artificial Constitution of *Genus* and *Species*' for the 'real Constitution of things'.⁹³

It should not be thought that in this attack Locke was simply appealing to his empiricist, abstractionist theory of universal thought as something independently assumed. More fundamental to his argument is his denial of real universals and appeal to the intuitive ontological principle that everything that exists is particular. That, if accepted, is enough to refute Aristotelianism, and his account of abstract ideas functions less as a presupposed premise than as a theoretical alternative to realism, or even as an explanation of the Aristotelians' mistake: in their theory of specific forms, genera, and essences, they take distinctions of thought for real distinctions, abstractions for realities. It is important here to grasp the general character of the liaison between anti-realist abstractionism and philosophical mechanism in the seventeenth century. For that purpose it may be helpful to consider a possible criticism of the anti-realists' strategy, to the effect that they had no right to their claim to have replaced real universals by similitudes arbitrarily picked out by the mind. This criticism need not rest on the blank logical ground that resemblance must be in some general respect, but arises because the new metaphysics seems to have given a universal essence or nature to matter as such. That transformation of Aristotelian natureless matter is the meaning of the doctrine that matter is itself a substance. Any answer to the criticism must therefore explain why the proposed essence of matter should have an ontological status any different from the allegedly inadmissible status attributed to Aristotelian essences of species.

The difference lies in what might be called either the 'actuality' or the 'perspicuity' of matter's essential attribute. Aristotelian essences are 'occult', consisting in powers or functions such as, in the case of man, rationality. We have to conceive of them, that is to say, through an intermittent actualization, virtually as the law or tendency governing the behaviour of the kind. That is why they were regarded as non-explanatory. Mechanism on the other hand, promised to explain all universal law and dark potentiality as mathematically derived from independently grasped attributes which are totally actual and directly perspicuous to the mind.⁹⁴ The paradigm for such an attribute is a thing's determinate extension.⁹⁵ Thus, whereas on the Aristotelian story the nature that two human beings or two pieces of gold have in common is irreducibly universal, for the mechanist, on the other hand, what makes two spheres or two cubes of matter behave in similar ways is not a common universal nature but a perspicuous resemblance. There is nothing in each case but extended substance within geometrically similar boundaries.⁹⁶ Just as corpuscularian mechanism could be supposed to give a lucid explanation of the substance-accident relation, so it provided a framework within which it might seem that universals could be reduced to resemblances.

There is, however, another way in which Locke's mechanist ontology lies behind his logic and theory of classification. The mechanist's world is one in which all differences are differences of degree, and everything, unless an atom, is in principle indefinitely mutable. For all differences and changes are ultimately just differences and changes in the spatial quantity and ordering and motion of the parts of things. Crudely, the particular complex perceptible things in existence, particular men, horses, oak trees, quantities of gold etc., constitute a vast plurality of machines among which there may be natural

structural resemblances, but no two of which, for all we know, are precisely alike.⁹⁷ How we should rank them on the basis of our observational knowledge is a matter to be more or less pragmatically determined. We should do so in a way which marks differences that are important to us and which at the same time fits our language for the purposes of communication.

Locke took the dual function of the Aristotelian specific essence to be, first, that of determining the boundary of the species by being present in, and only in, its members; and second, that of explaining or giving rise to the 'properties' of the species. The presence of the essence or form Locke took to be, for the Aristotelian, an all-or-nothing business; hence the supposed boundaries of the species are 'precise'. Locke did not quarrel with such a notion of precision. For, as we have seen, it is an extremely important and explicit principle of his philosophy (as of Frege's) that classes with precise, principled boundaries are a prerequisite for universal knowledge.⁹⁸ He held that what sets a boundary to the class is always what he calls the 'nominal essence', i.e. the abstract idea that embodies our criteria for the application of the kind-name or 'sortal'. What explains the properties of the species so defined, on the other hand, is corpuscularian structure (or, at least, something like it, if Boyle's theory is less than the whole truth). Those aspects of the structure of the individual members of a species which they have in common and in virtue of which they all possess the defining properties of the species, comprise what Locke called the 'real essence' of the species. The distinction between nominal and real essence derived, of course, from the Aristotelian distinction between nominal and real definition.

It is tempting to conclude from all this that Locke did not deny that each of the duties purportedly performed by Aristotelian essences is performed by something, but that he simply divided the labour between his two 'essences'. Yet that characterization of his position, however beguiling, is likely to mislead. For Locke really believed that nothing on earth could possibly perform the function that the Aristotelians ascribed to their specific essences or forms. Although the Aristotelian essence and Locke's nominal essence both define the boundary of the species, the former does so ontologically. If it also does so, for those who know it, epistemologically, as a criterion, that is quite accidental. But the Lockean nominal essence is intrinsically an epistemological essence and nothing more, a criterion by reference to which we mark off the members of the species. The boundary marked is a precise one which owes its existence to our drawing it: reality itself simply could not, in Locke's view, supply such a boundary. Reality can supply resemblances, but resemblances do not constitute natural boundaries. Resemblances do not draw lines.

The foundations of our system of classification, then, are not universal 'Forms or Molds, wherein all natural Things, that exist, are cast, and do equally partake', but the objective resemblances between things, the contingent fact that 'Nature in the Production of Things, makes several of them alike'.⁹⁹ This fact Locke regarded as undeniable and 'obvious' at the level of observation, and he accepted too (perhaps too readily) that phenomenal resemblances are an indication of underlying structural resemblances.¹⁰⁰ It is important to realize, as critics have often failed to realize, that Locke's recognition of natural resemblances is not a concession of any kind in the argument against natural

species, against natural boundaries independent of our concepts. It is true that it is on the basis of these observed resemblances that we form abstract ideas ‘and set them up in the mind, with Names annexed to them, as Patterns, or Forms, (for in that sense the word Form has a very proper signification)’. But the system of ‘species’ thus conceived is a necessarily inadequate system imposed by us on natural anarchy or continuity. Complex machines may differ from one another in indefinitely many ways, and by indefinitely small amounts or degrees. However we divide the biological or chemical worlds, there will be living creatures and stuffs with attributes that cut across the classes we have formed. Thus Locke can find the notion of a great chain of being, which was traditionally conceived of as a hierarchical order of distinct species, itself consonant with the anarchic mechanist vision:

There are Animals so near of kin both to Birds and Beasts, that they are in the middle between both.... There are some Brutes, that seem to have as much Knowledge and Reason, as some that are called Men...and so on till we come to the lowest and the most inorganic parts of Matter, we shall find everywhere, that the several *Species* are linked together, and differ but in almost insensible degrees. ¹⁰¹

Just as a species—i.e. on Locke’s view, a class of structurally and phenomenally similar individuals—may always crop up with some of the properties from one genus, and some from another, so may particular individuals with respect to species. And wherever we stop, we might always have made more divisions.

The absence of any natural basis, in the form of *boundaries*, for the classification of individuals existing at the same time is matched by anarchy over time. It should be remembered that the Aristotelian essence determines ontologically not only the boundaries of the species, but *ipso facto* the boundaries of the particular individual. The particular substance actually consists in, and is identical with, the specific essence or form embodied in its particular matter, to which the form gives definition. There is no question of the particular substance changing kinds. Change in the individual is change in its ‘accidents’. Any more profound change involves the replacement of one specific form by another, and so the destruction of that individual—hence the distinction in the Aristotelian system between ‘alteration’ and ‘substantial change’, i.e. substantial corruption and generation. This doctrine, as we shall see, is vulnerable to criticism, but it is not just an archaic quirk. On our ordinary, present-day notion of a natural kind or species there does seem to be some sort of ‘logical’ barrier to an individual thing’s changing its kind, and no barrier to its having some natural attributes other than those closely associated with, or explained by, its kind. It is thus a highly significant feature of the Lockean picture that each individual ‘machine’ is regarded as indefinitely mutable, at least in principle. Whatever its structure or constitution, it ‘may be changed all as easily, by that hand that made it, as the internal frame of a watch’. ¹⁰² Such change often occurs in the natural course of events. That is why the chapter ‘Of General Terms’ ends as it does. Its final paragraphs explain that essences are regarded as ‘ingenerable, and incorruptible’ only because they are abstract ideas ‘established in the Mind’. In the world

of particular substances,

All things, that exist, besides their Author, are liable to Change;

especially those Things we are acquainted with, and have ranked into Bands, under distinct Names or Ensigns. Thus that, which was Grass to Day, is to Morrow the Flesh of a Sheep: and within few days after becomes part of a Man: In all which, and the like Changes, 'tis evident, their real *Essence*, i.e. that Constitution, whereon the Properties of these several things depended, is destroy'd, and perishes with them. ¹⁰³

Thus if we say that quantities of matter or material things belong to kinds, we must admit that they constantly shift their kind, according as they satisfy now this and now that nominal essence. ¹⁰⁴ The real essence is not a universal substantial nature which, as it were, withdraws from the matter, nor is it a particular substantial form corresponding to an eternal essence in God's mind: it is simply those underlying structural features which no longer exist when the qualities constituting the nominal essence no longer exist. Such changes may be interpreted by us, from the point of view of our system of classification, as the generation and destruction of substances, but in reality nothing substantial is created or destroyed, just structure.

Now let us look at some of those features of Locke's argument which have caused particular difficulty for its interpretation. There are passages in which he might seem to have been saying, not that underlying ontological boundaries to species do not exist, but only that, if there are natural boundaries at the level of minute structure, they can have relevance neither to our actual scheme of classification, since we do not know them, nor to our methodology in natural history, since it is beyond our powers to discover them. Yet it is not difficult to see that some at least of his remarks upon which such a construction has been placed have been misinterpreted. First, to say as he did that members of a species defined by a nominal essence have, or probably have, corresponding similarities at real-essence level (so that we can talk of a corpuscular real essence of the species) is not to concede that the real essence could, independently of any human decision, determine ontologically the boundaries of the species. Locke was simply saying that, if we select some set of observable qualities to serve as our nominal essence, then no doubt behind this arbitrarily selected phenomenal resemblance will lie a structural resemblance indirectly picked out by the same arbitrary procedure of selection. He explicitly made the point that, even if we knew the 'real essence' in this sense, all the problems about boundaries would rise up again: 'For what is sufficient in the inward Contrivance to make a new *Species*?' Locke's standard example or analogy is the 'species' of watches whose 'inward contrivance' is known to watch-makers: 'But if any one will make minuter Divisions from Differences, that he knows in the internal frame of Watches; and to such precise complex *Ideas*, give Names...they will then be new *Species* to them, who have those *Ideas* with names to them'. The business of drawing boundaries and making divisions is arbitrary, at whichever level of difference it is conducted. The same goes for *man* as for *watch*: a normal man as evidently differs in structure ('the Wheels, or Springs...within') from a changeling as a changeling from a drill, but

'whether one, or both these differences be essential, or specific, is only to be known to us, by their agreement, or disagreement with the complex *Idea* that the name *Man* stands for'.¹⁰⁵

With all this spelt out (and we shall consider another argument of Locke's that crosses the *ts*) there is little question but that in the few places where Locke seems explicitly to have made the concession that natural boundaries might exist he did so at least primarily for the sake of argument. The only alternative, one not clearly supported by the context of these passages, would seem to be that he suddenly felt compelled by extreme sceptical doubts to allow that everything he had been talking about may after all be quite beyond our comprehension. That the first diagnosis is at least reasonable is evident from those passages in which he carefully isolated 'the usual supposition' of *Aristotelian* essences, in order to prove that, *even if such essences existed*, the boundaries they set to species could not be known.¹⁰⁶ To read the 'concession' in that argument as the concession of a genuine possibility, when it was offered virtually in the same breath as allegedly conclusive ontological refutations of Aristotelian doctrine, would clearly be absurd.

The subtle epistemological argument against substantial forms is worth examining. Locke argued not simply that, since we know nothing of such forms, our ordinary classification must be on another basis, but that, if a system of distinctions along the borderlines marked out by Aristotelian real essences were to get off the ground, these real essences would have to be known; and yet they could not possibly be known. It is important to realize that, as Locke himself would have been well aware, the first lemma of this latter argument needed proof just because it was no part of the Aristotelian case that we normally refer to the real essence as a criterion of application for the term. On the contrary, except in the case of artifacts,¹⁰⁷ it was commonly supposed that we normally do not know the specific essence, although of course if we do know it the task of classification may be made easier. How then are we supposed to distinguish members of the species? It was thought that we could identify the 'properties' of the species by induction from individuals and then, by philosophical reflection, pick out the peculiar 'difference'. The purpose of identifying the real essence, as we have seen, was explanation, not classification: it was simply and not unreasonably presupposed that we can identify members of the same species at least fairly well on the basis of more or less obvious points of resemblance recurring in our experience. That is why a distinction between nominal and real definition was already built into the theory that Locke was attacking. A nominal definition, as was explained in chapter 2 above, consists of a non-explanatory list of attributes peculiar to the species and so suitable as criteria of recognition.

Locke, of course, was ready to argue that any specified candidate for the role of Aristotelian real essence would turn out to be a non-explanatory nominal essence, and that the explanatory role in respect of the associated 'properties' could be fulfilled only by corpuscular structure or something like it.¹⁰⁸ But his purpose within the epistemological argument was to attack the assumption that we could even begin to discuss the question of the (Aristotelian) real essence of 'this species' which we have already identified through its properties or from examples. How would we know which attributes of the particular things or of any single thing before us were *properties* rather

than *accidents*, and therefore 'so annexed to it, that any of them being away, we may certainly conclude, that that Essence is not there, and so the thing is not of that *Species*'? ¹⁰⁹ Locke agreed that if we knew either the properties or the essence, we should be able to distinguish members of the species: but we could not know the properties unless we knew the essence. It is really the boundary between property and accident that is in question in his critique of the Aristotelian notion of how we identify species, and his demand was for the Aristotelian to explain this boundary, and how it could be known. Unless it is known, no species has been identified, and so the question of what 'its' real essence is cannot arise. If it had been said that we can identify a species by means of an ordinary word such as 'gold', Locke's response would have been that this is either to make use of an existing nominal essence in his sense, which imposes a boundary on reality for us, ¹¹⁰ or (worse still) to assume that there is some precise boundary marked out by a mere word in general use, i.e. by that loose and imprecise usage without clear and distinct ideas which he called the 'civil' use of language, and which he contrasted with the precise or 'philosophical' use necessary for universal science or the systematic study of anything at all. ¹¹¹

If, on the other hand, it had been said that we can identify the species by means of a single individual, as the species, whatever it is, to which *that* thing belongs, Locke would have been ready to challenge the Aristotelian to apply the distinction between properties and accidents to the individual. The individual member of a species cannot by itself determine the boundary, since 'there is scarce any particular thing existing, which in some of its simple *Ideas* does not communicate with a greater, and in others with a less number of particular Beings'. ¹¹² 'No one has Authority to determine the signification of the Word *Gold*, (as referr'd to [that sort of Body the Ring on his Finger is made of]) more to one Collection of *Ideas* to be found in that Body, than to another.' ¹¹³ It is true that a corpuscularian sense can be given to the notion of the 'real essence' of a particular individual—in a way, Locke suggested, a sense that more closely accords with the original meaning of the word 'essence'. In that sense the 'real essence' is the 'very being of any thing, whereby it is, what it is', the underlying 'internal constitution' of the particular. ¹¹⁴ But 'it will be found a quite different thing, to argue about Gold in name, and about a parcel of the Body it self, v.g. a piece of Leaf-Gold laid before us'. ¹¹⁵ Locke was rejecting the notion that an individual belongs to one and only one kind: that is, he rejected the distinction between generic and specific names, and the notion of an ultimate or last species, the species to which uniquely the individual is indissolubly tied. For if we consider the 'real essence' to be the underlying structure of the particular, without reference to a name, it has to be considered as a whole: and then it will appear that all the qualities and natural attributes of the thing, whether classed by us as 'difference', 'properties', or 'accidents', flow equally from the 'real essence'. Indeed, 'particular Beings, considered barely in themselves, will be found to have all their Qualities equally *essential*; and every thing, in each Individual, will be *essential* to it, or, which is more true, nothing at all'. ¹¹⁶

Thus Locke was prepared to direct more or less the same criticism against the suggestion that there might be real corpuscularian boundaries to species as he used against the Aristotelians, i.e. that the property-accident distinction can be applied only

relatively, in relation to a nominal essence. This central point is made clearly and emphatically at the outset of the chapter on the names of substances, where the focus is brought to bear on changes within the individual. First, there is a famous appeal to experience in support of the contention that no observable attribute is necessary to the particular:

take but away the abstract *Ideas*, by which we sort Individuals... and then the thought of anything *essential* to any of them, instantly vanishes.... 'Tis necessary for me to be as I am; GOD and Nature has made me so: But there is nothing I have, is essential to me. An Accident, or Disease, may very much alter my Colour, or Shape; a Fever, or Fall, may take away my Reason, or Memory, or both; and an Apoplexy leave neither Sense, nor Understanding, no nor Life.
117

Now he turned to the real essence, 'that particular constitution, which every Thing has within it self, without any relation to any thing without it', i.e. the intrinsic or 'internal' structure which is the foundation of its properties. Yet 'particular' Locke implied, must here be construed as *specific*, and 'thing', as *sort of thing*:

But *Essence*, even in this sense, *relates to a Sort*, and supposes a *Species*: For being that real Constitution, on which the Properties depend, it necessarily supposes a sort of Things, Properties belonging only to *Species*, and not to Individuals.¹¹⁸

Locke's example is again gold, defined by colour, weight, malleability and fusibility. The real essence is the structural foundation of 'these Qualities, and their Union', while whatever other qualities this underlying constitution regularly gives rise to, such as solubility in aqua regia, are properties. 'Here are *Essences* and *Properties*, but all upon supposition of a Sort, or general abstract Idea...considered as immutable: but there is no individual parcel of Matter, to which any of these Qualities are so annexed, as to be *essential* to it, or inseparable from it.' In other words, real essences could not determine species ontologically because real essences are relative to nominal essences. The predicables do not apply at the level of species *de re*, but only *de dicto*.

These arguments, which because of their importance have been presented at some length, are perfectly clear and, moreover, plausible. The argument of III.iii.4 appeals to experience, but that of III.iii.6, which does not, shows that such an appeal is unnecessary. For, on the corpuscularian hypothesis, both complex individual objects (think of them as discrete, unified machines) and parcels of stuff (i.e. matter) are in principle indefinitely mutable structurally, and therefore perceptibly. An object which boringly retains a set of determinate attributes throughout its existence does *so per accidens*.

Since these arguments are the logical corollaries of Locke's vision of a world in flux, let us look at a line of his thought which is sometimes interpreted in a sense hardly compatible with them. First, suppose that someone makes something which he claims is

gold. It is like gold in all obvious, perhaps even in all known respects. Yet suddenly it is found that in one, perhaps startling, respect it is unlike 'natural' gold. Apparently one such test (called 'cupellation') was discovered in early times and afterwards regularly used to unmask pretensions to aurification. Yet, centuries before Locke, some had argued that, since the defining qualities were all present, gold had indeed been made in such a case: slightly peculiar gold, but not mere imitation gold. Such an argument was never very plausible,¹¹⁹ and we have for a very long time thought of gold and other substances in such a way that two parcels of pure gold must be presumed to have all natural attributes in common apart from shape, being hot or cold, being liquid or solid, and so on. 'Accidents' are cut to a minimum.

Locke's official answer to the question about the imitation gold would have been that, unless the cupellation test had been included in the nominal essence or definition of 'gold', it is gold.¹²⁰ Yet he was clearly aware of the unorthodoxy, even paradoxicality of that answer. Moreover, he was well aware that systematic experimental evidence of differences between instances taken to be of the same stuff would show not merely that the accepted classification was not based on known scholastic essences, but that it needed improvement. Indeed, he consciously associated himself with the positive programme of Boyle and the Royal Society for improving classification: things should be sorted, as Boyle put it, 'as they deserve'.¹²¹ The question is how both Locke and Boyle could have thought that things may deserve to be sorted one way rather than another, and yet deny that there is a right way to sort them determined by objective boundaries between kinds.

The answer to this question is readily to be found in the *Essay*, in the elaborate explanation of how and why we should improve and remodel our definitions in the light of experience, an explanation which self-consciously avoids the implication of natural boundaries. The topic of the 'rectification' of our nominal essences is treated under the headings of 'clear and distinct ideas' and the imperfection of words. These discussions have to do with a method for raising language above the confusions of its 'civil' use, so as to give it scientific or 'philosophical' precision and consistency.¹²² Men too often 'content themselves with some few obvious, and outward appearances of Things, thereby readily to distinguish and sort them for the common Affairs of Life'. Biological species are distinguished in some men's minds by not much more than shape, and chemical species, by colour.¹²³ Indeed we commonly think about kinds of substances so vaguely that sortal names have no precise, settled meaning for us at all. We pick up a variety of such names, supposing them each to denote a distinct kind of thing, without appropriately distinct ideas.¹²⁴ Such confusion, so Locke continually emphasized, both promotes and is compounded by the doctrine of real species named by the sortals in our language.¹²⁵ We lazily and thoughtlessly suppose that there is such a real species as 'liquor', for example, and argue about it without ever getting clearly before our minds just what is to count as liquor and what is not.¹²⁶ What is called for, Locke thought, is a combination of decision and 'natural history'. We need to recognize *both* that there is always something arbitrary about the choice of the nominal essence *and* that it can be done well or badly, depending 'upon the various Care, Industry, or Fancy of him that makes it'.¹²⁷ It is for this purpose, the avoidance of confusion, that Locke proposed his Natural History dictionary: to remedy the 'mistake, that the signification of common Words, are certainly

established, and the precise *Ideas*, they stand for, perfectly known; and that it is a shame to be ignorant of them'.¹²⁸

In short, Locke believed that the chief trouble with the classification of his time was its sloppiness, and that one of the chief barriers to its improvement was an uncritical attitude to ordinary language. His remedy was a programme of agreed, precise definitions based on careful observation and experiment, adequately distinguishing substances that we find we need to treat as distinct. Only for this reason did he agree that 'to define their Names right, natural History is to be enquired into'.¹²⁹ He was not endorsing the view that classification involving even such 'clear and distinct ideas' could well suffer from the fault of not distinguishing what more observation would prove to be ontologically distinct species. To interpret him as if he were is to import into his argument contradictions and tensions that do not exist there. Indeed, the whole argument was aimed at refuting this view and explaining it away as an error natural enough for men to make, 'especially such as have been bred up in the Learning taught in this part of the World'.¹³⁰ Mere observation of differences could at worst show that two parcels of 'X' have as much right to be treated as of different species as a parcel of 'X' and a parcel of 'Y'.¹³¹

True, in the course of this argument, Locke went so far as to say that,

if the formal Constitution of this shining, heavy, ductil Thing (from whence all these its Properties flow) lay open to our Senses, as the formal Constitution, or Essence of a Triangle does, the signification of the word Gold, might as easily be ascertained, as that of Triangle.¹³²

Yet this means, not that we would then perceive the objective boundary of a species, but that we could then fix and agree on a nominal essence consisting of a relatively few mechanical properties, as in the classification of machines with observable working parts, or indeed of geometrical figures: there would not be the same room for confusion as there now is, when we have to select defining properties from an indefinitely large number of powers. So too his discussion of the inadequacy of our ideas of substances (which contain neither ideas of all the observable properties of the substance nor ideas of the structural attributes that constitute the real essence) is irrelevant, apart from the passing digs at the Aristotelians, to the question of naturally bounded species. For he was there concerned with the inadequacy of our ideas of what falls within the boundaries that we impose on reality.¹³³

How is Locke's argument to be judged? Given an existing system of classification, in both biology and chemistry, which was imprecise, arbitrary and, on the evidence of careful experiment, massively coarse, given the absence of any clear hypothesis, such as the hypothesis of identifiable elements and chemical combination, which could serve as the basis for an alternative system, and given the genuine difficulty of combining the notion of precise specific boundaries with the assumption of comprehensive mechanical laws governing corpuscularian matter or something like it, his approach to the problems of classification, coherent as it is with a systematic epistemology and theory of meaning, is a shining example of philosophical intelligence. In some respects we might compare his situation with that of a sociologist who feels that his discipline makes too much

uncritical use of imprecise and undefined jargon. Such a terminological reformer might reasonably hesitate to propose a programme of discovering and carefully naming supposedly objective sociological 'kinds', as if the phenomena he studies were neatly packaged by reality. Nor could he seriously urge the general acceptance of a body of precise, but quite arbitrary, armchair definitions. His proposals must contain, as Locke's did, all three elements of general agreement, reference to reality, and motivated selection and decision.

Are there real species?

What is wrong with Locke's account of the names of substances? Although he allowed that where the observable attributes constituting the nominal essence of, say, 'water' recur, some underlying structure will reasonably be supposed also to recur, he nevertheless insisted that that supposition is semantically inert. For this he gave, broadly speaking, two reasons. First, it is impossible that our employment of the term should be determined by what is unknown to us; and, second, the supposition of a real essence can supply no boundary to the species other than the boundary supplied by a nominal essence, i.e. an arbitrary definition. These reasons may be connected, but they raise rather different issues.

Take first Locke's belief that, because the supposition of recurrent aspects of otherwise unknown structure inertly follows the observation of recurrent qualities and powers (the two cannot come apart), all the semantic work must therefore be done by those observable qualities and powers. Leibniz objected that our criteria for the application of a term may constitute not so much a strict definition, as a catalogue of indications of an unknown underlying common nature. As Saul Kripke and Hilary Putnam have more recently emphasized,¹³⁴ the 'nominal essence' associated with a name may after all come apart from the real essence of what the name names. Different people may employ different criteria for the application of the word 'gold' and yet mean the same by it just because they use it for the same stuff; while, on the other hand, the same stuff may appear in different guises, so that it may be present even when a standard criterion of application is unsatisfied. One illustration of the latter possibility is supplied by chemical composition: copper may be present (and recognizably present) in copper oxide or copper sulphate, when none of the usual indications of elemental copper is present. For a word like 'gold' or 'copper' or 'horse' is characteristically introduced into language demonstratively or ostensively, rather than through any sort of rigid definition: the stuff or thing is, as it were, christened.

That the presumption of an underlying nature does after all have a semantic role is also illustrated by another kind of case. Suppose that, misled by their phenomenal resemblance, people have always confounded two structurally distinct, equally common substances, calling both 'salt'. On the anti-Lockean view there are, broadly speaking, two possibilities. Very roughly, if (despite the differences) there is a fundamental generic resemblance between the two substances, then 'salt' is the name of a genus however much its users may believe that it is the name of an ultimate species. 'Salt' is not equivocal, but there are two kinds of salt. If, on the other hand, there is no significant fundamental or underlying resemblance, then 'salt' is at best ambiguous despite its users' belief that it is univocal. Their beliefs about 'salt' are like beliefs about 'Cato' held by

someone who has conflated Cato the elder with Cato the younger. It will be seen that whether the (general or proper) name is ambiguous will depend on the world, and the relationship of the world to the speaker, rather than on the speaker's beliefs or 'concepts'.

Now suppose this second state of affairs, and then suppose that it is discovered that the presumption of a single common nature is false. What will happen to the word 'salt'? First, in recognition of the ambiguity, it may simply earn two entries in the dictionary; or, perhaps, for convenience it may come to be employed for just one of the substances, a new term being coined for the other. Alternatively it may for some pragmatic reason continue to be employed for both substances as a univocal word, according to a definite criterion. In this last case it has become a quite different kind of word. From being a problematic name of a natural kind it has become a grossly descriptive, possibly even functional term, comparable to 'powder', 'seasoning' or 'sedative'. It is now, as it was not before, more or less rigidly associated with a certain criterion of application. It is therefore not a sortal term, such as classifies things according to their common nature, nor is it even a would-be sortal.

There is in fact a certain tension or inconsistency in Locke's account. Although he was unyielding in his belief that the presumption of an unknown structure can have no semantic relevance once the sortal term has been introduced and defined, he nevertheless allowed the rationality of such a presumption to play a role at an earlier stage, guiding us in the *formation* of substance-ideas. In general there is the implicit suggestion that repeated observation of coexisting qualities and powers is requisite to justify the presumption of a recurrent underlying cause of their union, and so to justify the formation of a complex idea. More particularly and explicitly, he remarked:

Whosoever first light on a parcel of that sort of Substance, we denote by the word *Gold*, could not rationally take the Bulk and Figure he observed in that lump, to depend on its real Essence, or internal Constitution. Therefore those never went into his *Idea* of that Species of Body. ¹³⁵

The point must be that the gross size and shape of a lump of gold, unlike that of a horse, are readily seen to be indefinitely mutable without consequence for the other distinctive attributes which it possesses. They cannot therefore be supposed to share a common basis with those attributes. The only specific constitution which can be presumed is that of a homoeomerous substance or stuff. In at least two ways, then, Locke suggested that reasonable beliefs about the functioning of the internal constitution may come before any nominal essence has been formed, as a condition of its being formed legitimately or rationally.

Yet, although Locke believed that there are proper and improper ways of forming ideas of substances (a belief given formal expression in the distinction between 'real' and 'fantastical' ideas) it seems that he found nothing semantically ill-formed about a 'fantastical' idea once the rule against forming such an idea has been breached. Its name can make the same contribution to a proposition as the name of any 'real' idea: 'twill be altogether as true a Proposition, to say *all Centaurs are Animals*, as that *all Men are Animals*.... For in both Propositions, the Words are put together according to the

agreement of the *Ideas* in our Minds'.¹³⁶ On the anti-Lockean view, on the other hand, 'All men are animals' is not in any case a mere truth by stipulated definition, while simply to invent such a term as 'centaur' is to fail to introduce it into the language at all, at least as an unproblematically meaningful name. There are no exemplars with which the name can be ostensively linked, and no deep resemblance underlying the criteria of recognition. It may in a sense be true that centaurs have hooves, but they have no genes, their life-process is neither carbon-based nor not carbon-based, and we cannot sensibly ask whether the centaurs in this picture could interbreed with the centaurs in that story (or whether 'centaur' is the name of a species or a genus). Consequently the term 'centaur' has no determinate satisfaction-conditions: or rather, in so far as it was introduced without reference or regard to anything real, it could not possibly be satisfied by anything real. To know the imaginative origin of the term is to know that it is not the name of a real kind. It is therefore to know that, whatever strange creatures exist on Mars, they are not centaurs. Like 'Mr Pickwick', the name 'centaur' is a semantic cripple, incapable of venturing far away from particular intentional contexts, and hobbling even there. Locke's view, however, was that 'centaur' has determinate satisfaction-conditions no different in kind from those of 'horse', conditions provided in each case by a rigid nominal definition.

Our ignorance of real essences is not the only ground of Locke's position. A second ground explicit in his argument is that only human definitions or concepts can supply specific boundary-principles, since real essences are themselves distinguished only relatively to nominal essences; and that, even with full knowledge, arbitrary definition would be necessary. Now the problem of boundary-principles, oddly enough, is one that is not directly addressed by the modern anti-Lockean argument. In general, proponents of that argument have seemed happy to assume that a natural-kind predicate 'alpha' can always be introduced into a language through ostension combined with a formula such as 'Whatever animal is of the same kind as this one is an alpha'.¹³⁷ Yet the possibility of such a simple ostensive procedure is in no way entailed by the point that deep common resemblance is a semantically significant condition of satisfaction of a sortal predicate. As Locke argued, mere resemblance draws no boundaries, and what superficial resemblance cannot do, deep resemblance cannot do either. To talk of *relevant* resemblance would be to beg the question, leaving it unexplained what principle of relevance to species-membership there is other than an arbitrary definition. What this simple conception of ostensively definable kinds requires is not just that deep resemblance should be semantically significant in the ways that Kripke and Putnam have proposed, but that there should actually be natural boundaries, i.e. that there should be, in Locke's words, 'chasms or gaps' in nature. The problem therefore remains for the adherent of the modern view: are there such natural boundaries? If so, in what do they consist? If not, in what could they have consisted?

Because of the emphasis placed by the modern logical doctrine on ostension or deixis, it may seem surprising that doubt can be cast on the efficacy of the ostensive formula (and of the appeal to the relation, *of the same kind as*) without its also being cast on the general principles of the modern approach. That this is so, however, can be seen from the notorious difficulties in the notion of a wholly objective or 'natural' biological taxonomy.

This is hardly the place for a comprehensive discussion of taxonomic theory, but it may be helpful to say something about its relevance to semantics.

In a classic and witty critique of the assumption that the division of plants into families and genera corresponds to natural groups, S.M. Walters in 1961 considered the evolutionary hypothesis that the larger groups at each level are those that have been longer in existence. This principle is true, he argued, provided that the age of the group is understood with reference to the history of taxonomy rather than, as the author of the hypothesis had intended, as evolutionary age. Angiosperm taxonomy has been massively influenced by the accident that it began with common distinctions made among those European plants which came to be used as exemplars for the division of plants from elsewhere. Moreover the grounds of the early distinctions were often more or less extraneous. Thus 'the economic importance of the grasses in Europe had ensured that by the time of Linnaeus forty-six genera were named and described, as against five of the present-day *Cyperaceae* (sedges)'. Linnaeus permissively accepted the existence of a utricle round the nut as the mark of a single genus of the latter, with the result that the genus *carex* contains 'well over a thousand species in the world'. In general there are still about three times as many species to the genus among sedges as among grasses. The ostensive formula, 'Whatever is of the same genus as this is an alpha', could only cause a smile among self-aware taxonomists.

Walters' conclusions, with their combination of scepticism, pragmatism and conservatism, have an astonishingly Lockean tone:

Any attempt to improve...classification should logically be preceded by a statement of the *purpose* of the classification... [which] seems clear enough; it gives a general map of the diversity of living organisms, and creates groups about which it is possible to make a wide variety of inductive generalizations.

Here I am driven to plead for conservatism. Could we not admit frankly that the classification we have, and are obliged to use, is different from the one which we might all be using if we could wipe the slate clean and start again, but that it nevertheless serves reasonably well.... Further, can we not all agree that the piecemeal 'improvement' of that classification as each piece of new information (or what is worse, new speculation) becomes available, brings in its train practical disadvantages in the form of changes to familiar names which usually far outweigh any advantages which result from the incorporation of new information? This need not mean that *no* change is ever made; but it should mean that the practical consequences of such a change should be carefully weighed before a recommendation is made....

Does this reduce the theory and practice of taxonomy to the mere provision of a convenient reference system for botany? The answer is that it does; but it is not necessary to use the pejorative verb 'reduce' and the adjective 'mere' to describe the aim of taxonomy.¹³⁸

With his emphasis on morphology and the practical purpose of classification, Walters seems to have come close to the pure Lockean conception that an arbitrary nominal

definition rigidly determines membership of the genus. Nevertheless he does not deny outright that taxonomists should in general employ such systematic morphological resemblances as they presume are due to deeper, more explanatory connections. His objections to revision rest rather on his sense that botanists' ideas about deeper connections are speculative and forever changing, so that too great a deference to those ideas produces an inconveniently variable system of classification. But it is worth setting against this scepticism the classical evolutionary ideal, here expressed by E.B.Ford:

It is necessary to classify any large assemblage which we wish to study so as to reduce it to order and to divide it into a series of groups which may be identified by some logical system. This is true whether it be composed of butterflies, words, or the figures of heraldry. There are usually many ways in which a classification can be made, but that adopted for animals is based upon [genetic] relationship...for it [relationship] reflects a wider and more important aspect of reality, the course of evolution, than does an arbitrary arrangement made for some particular purpose. Indeed, since it does reflect this wider aspect, it will undoubtedly be, in general, the most satisfactory system. Moreover, when expressed as a diagram, it becomes a genealogical tree and, if it be successful, we can make a generalization of the utmost importance about it: that all the animals in any group, whether it be a great one like the insects or a small one like our common White Butterflies (the genus *Pieris*), are more closely related to one another by actual descent than they are to any other organisms upon earth.¹³⁹

We may well be struck, with Walters, by the difficulty of knowing when biological classification is by this high standard 'successful'. Yet even Ford's ideal does not envisage the elimination of all human choice. As he went on to say, 'the scope of these groups is necessarily somewhat arbitrary, [although] the most natural of them is the species'. Moreover we cannot suppose that this arbitrariness is restricted to the determination of the degree of cognateness required for membership of the same taxon at each level. First, it is difficult to imagine a clear, universally applicable principle of measurement of evolutionary relationship; and second, even if there were such a means of measurement, it could not by itself determine the boundaries between taxa. If new taxa arise in the course of evolution, then it follows that the same general relationship which exists within the taxon must be capable of straddling a boundary between taxa. Nor will a move, recently made technically possible, to the direct consideration of DNA solve the difficulty for the hypothesis of natural taxa. Physically measurable differences in genetic constitution need to be complemented by a judgement of their evolutionary significance. What, perhaps, can be done is to employ morphological and genetic parameters in order to plot existing species, with an eye too on the fossil evidence. If species then fall conveniently into bunches, primary bunches can be treated as genera—provided of course that biologists are prepared to suffer the disorientation involved in more or less 'wiping the slate clean'. Any outlying species could be given similar status, on the grounds that they are incipient or potential genera, distinguished from their closest relations by the

sheer extent of their mutation. If the bunches bunch, there are natural families too.

If such an ordered result is not forthcoming, taxonomy will still be needed for practical purposes. We shall doubtless continue with 'arbitrary' divisions of plants and animals into families and genera, even if there is no possibility of an objective answer to the question of how many there are of each. That is not to say that 'nominal essences' would have just the role that Locke ascribed to them. Unknown deep connections could still contribute to the satisfaction-conditions of taxon-names. That is to say, new discoveries at the deeper level could lead us to recognize that names had been misapplied. In that case, taxon-names would still be names of 'natural kinds' in the sense of the modern theory. However uniquely characteristic of birds feathers are taken to be, there might still, in principle, turn out to be featherless birds or feathered reptiles. The semantic role of deep connections might be inhibited by pragmatic conservatism without being eliminated.

To see evolutionary relationship as the reality to which an objective system of classification should aim to approximate, and as a constitutive element in the 'deep connection' between members of a kind, is of course to live in a different world from that of Aristotle or Locke. Both philosophers take it that the underlying or 'real' essence lies in the intrinsic constitution of members of the kind. This difference, between a relational and a constitutional view of what unites its members, is of no little importance if we turn to the question so far unconsidered, 'Are there real species?' Origin or descent is for Aristotle something which goes along with the specific essence (according to his principle that 'man begets man'), but it is hardly a condition of membership in its own right, and can have nothing to do with the grouping of species into genera. For Locke too origin is at best an indirect, insufficient and not very useful criterion for deciding on species membership: 'must I go to the *Indies* to see the Sire and Dam of the one, and the Plant from which the Seed was gather'd, that produced the other, to know whether this be a Tiger or that Tea?' ¹⁴⁰

Oddly enough, both Kripke and Putnam seem to agree with the old-fashioned constitutional view of kind-membership. They discuss biological kinds in just the same terms as chemical kinds, with membership of the kind supposedly determined by 'internal structure'. But if we are to look for a natural boundary round the species, internal structure will not supply it on its own. As Locke rightly remarked, as far as the 'wheels and springs within' are concerned, a 'changeling' may not unreasonably be supposed to be as different from a normal man as the latter is from a 'drill'. ¹⁴¹ Employing any straightforward or natural method of quantitative physical measurement, the difference between the genes of a chimpanzee and those of a man may be smaller than the difference between the genes of a normal man and those of a man suffering from a genetic abnormality. That is not to say that we could not construct a 'real definition' of man in terms relating solely to DNA, i.e. a description of genes (presumably disjunctive) in purely physical terms which includes the changeling and excludes the chimpanzee. Such a 'definition' could take into account at least known genetic variations and deformities. But the point is that a purely constitutional definition or description of a 'gene-pool' presupposes, and is answerable to, another kind of boundary-principle, a principle which explains why the difference between a sufferer from Down's Syndrome

and a normal human being is insignificant by comparison with the difference between a normal human being and a chimpanzee; and indeed is absolutely insignificant from the point of view of natural classification.

Locke thought, in effect, that the need for such a further principle means that specific boundaries are conceptually imposed, which is why he laid so much emphasis on the nominal essence. Yet nature does supply another sort of boundary in all that is involved in the members of one species forming a historically related, unitary group. That is a matter, if sometimes a rather loose one, of the interbreeding and common origin of members of the group. Hence the popular view among philosophers of taxonomy that species are historic individuals. On this view, if we somehow synthesized a quasi-animal, however accurately the morphology and genetic structure of *Tyrannosaurus rex* were reproduced, the result would still not be a member of that species, or of any species.

It is of course no accident that members of the interbreeding group share peculiar similarities of macrostructure and microstructure as well as a common origin. The existence of the group depends on such similarities. But reciprocally the significance of just these structural similarities by comparison with others, perhaps as striking, is precisely that they make such a group possible. Without the group, they would have no special significance. They would just be resemblances. Thus the notion of a 'gene-pool' in fact supplies a hybrid or comprehensive principle which combines structural or constitutional considerations with considerations of a quite different type, namely relationship and the capacity to interbreed. It is only the latter (linked as they are to the former) which make it possible to introduce a term by pointing to a strange creature and saying 'Whatever is of the same species as that is an *alpha*'; or to suppose that to count species is not just to count human distinctions or names.

It will be seen from all this that if there are natural or real biological species, there might not have been. The kind of limitations on interbreeding which, together with the laws of heredity, make structure as trustworthy an indication of origin as it is, might conceivably not exist in some possible biological world. To imagine a fairly extreme case of such a world, we might first imagine that the characteristic differences that we associate with 'race' (differences generally due to the geographical isolation of different parts of the species) existed, but were not inherited; then, a world in which the same was true of all specific characteristics among, say, warm-blooded animals. The resultant chaos is worth considering with respect to the logical relationship between individual and species.

The Aristotelian essence, as it has been said, determines the boundaries of the individual as well as of the species. The individual exists only in so far, and as long, as it instantiates the specific essence. For Locke, on the other hand (at any rate, in the context of the argument about species), the individual can survive change of species just because it has in principle an indefinite capacity for surviving any mechanical change short of material disintegration. In rejecting the ontological distinction between properties and accidents, Locke asked, 'what are the alterations may, or may not be made in a *Horse*, or *Lead*, without making either of them to be of another Species?'¹⁴² It is that question, not the question of existence or identity, which he thought is 'easy to resolve' by reference to the nominal essence, but unanswerable by reference to the supposed real essence.

The ontology of modern science, and the framework it supplies for the relationship between biology and physics, is of course far closer to Locke's than it is to Aristotle's. As in general biological individuals fall naturally into species, each with a typical structure, so they normally, but not inevitably, undergo a typical life-history determined by that structure. But as even the original structure of a living thing can be atypical of its species, so can its life history. At the extreme, we can conceive of the possibility in principle that a horse should have its DNA so altered by some experiment in genetic engineering or by radiation (or whatever) that it grows to be in all respects like a deer. Given the nature of specific boundaries, the resultant creature does not have the right origins actually to *be* a deer. But there is no reason why it should not be a horse, just as there is no reason why a 'changeling' should not be a man. It is a horse which has suffered deformity, not a deer which has replaced a horse. So there is something wrong and something right both in the Aristotelian view and in Locke's. The Aristotelian recognized a tight link between the biological individual and its species: biological individuals do not change kinds. He found the explanation of this link in the theory that the individual cannot outlast its underlying constitutional specific essence or form, its intrinsic nature. Locke, on the other hand, saw that there is not some specific essence or structure in the individual which is in principle incapable of change, or which the individual cannot in principle outlast. Nor is there some sequence of changes to which the individual is in principle fated or restricted. He wrongly concluded that there is not a species to which it is restricted.

The rigid link between the individual thing and its original species is not an arbitrary linguistic rule, like some rule that birth determines caste, or that a man born in England can never be anything but an Englishman. It exists because of the typical life-history which normally springs from that origin and which, together with the normal structure, will serve as a criterion of species membership. Without those connections there would be no point to a system of classification determined by relationship or descent. Yet, equally, the rigid link between individual and species does not presuppose a similar rigid link between the individual and specific structure or specific life-history. Whatever structure or life-history is believed to be typical of a horse, it is possible to suppose, or with the resources of science perhaps one day to bring about, a horse which falls radically far from the norm. Not a fairy-tale 'horse', but a genuine member by origin of the earthly family of creatures to which we have given the name.

It follows that membership of its species is not what *determines* the boundary-conditions or existence-conditions of the individual. What then does supply these conditions? In what does the existence of the individual consist? In Book III of the *Essay* Locke adopted a simple, ancient view which deserves much more notice than it receives today. Boyle was more interested in chemical stuffs than in individuals, but he was expressing that view when he remarked 'that we need not deride the ancient atomists, for attempting to deduce the *generation* and *corruption* of bodies from the famed *synkrisis kai diakrisis*, the *convention* and *dissolution*...of their supposed atoms'.¹⁴³ The individual thing is a mass of particles functioning jointly. It has the boundaries and unity of a discrete, coherent, complex 'machine'. The relationship between its parts is not mere local conjunction but material coherence, which is a dynamic, lasting relation, not a momentary state. Material unity is itself a function of the jointly operating parts. That is

why the individual can *exist through* the loss of all the Aristotelian generic and specific differences of *man*: reason, sense and life itself.¹⁴⁴ Certain powers drop away from the individual, certain processes cease. The process of life is replaced by the process of decay. It is only when the individual (or enough of it) ceases to cohere as a discrete material unity that it no longer exists. Yet, *contra* Locke, as long as that individual exists, it remains a member of its species. Even a dead man is properly called a man, and the corpse, as much as the living man it was, owes its material unity to its origins. This model for the existence of complex substantial individuals will be explored in Part III, below, and set beside rival models, old and new. The comparison will not be to its disadvantage.

Something should be said about the natural kinds of stuff, the chemical species. Here Locke's penchant for precision seems no great fault. The natural boundaries of biological species are rough, and with respect to them it can seem a weakness in his argument that it assumed a more rigid, mathematical link between the properties and essences of substances even than the Aristotelian tradition always proposed. Aristotelian doctrine allowed, for example, that injury might prevent characteristic expression of the essence, yet when Locke continually recurred to the topic of monsters and idiots he took it as an agreed principle that 'it is as impossible that two Things, partaking of the same real *Essence*, should have different Properties as that two figures partaking in the same real *Essence* of a Circle should have different properties'.¹⁴⁵ This precise, geometrical ideal is intuitively more attractive (and has borne fruit) in chemistry just because chemistry is closer to basic physics. Origin counts for nothing, structure or constitution rigidly determines the kind. Apparent differences between uncombined members of the same kind are standardly explained as due to impurity and the like.¹⁴⁶ Our understanding of their physical nature allows some 'accidents', such as heat and cold, freezing and boiling and melting, and upholds the pre-theoretical judgement that such transitory structural states are irrelevant to the kind. The theory of elements and chemical combination has confirmed the ordinary or primitive view of natural kinds with independent sharp edges, refuting Locke's alternative picture of a chemical world with no 'chasms or gaps'. That picture was based in part on Boyle's critique of the earlier doctrine of elements, but Boyle's conception of the combination of particles to form a vast variety of stuffs itself provided some of the materials for the modern theory.

A great difference from biology appears too when we turn to the relation between species and individual, for the 'individual' in the sphere of chemical stuffs is a particular *quantity* of the stuff in question. Such quantities do not, as such, have 'normal' life-histories, and certainly cannot have 'abnormal' ones, but simply react to their circumstances. Yet the question of what changes a quantity of a chemical can survive is perhaps not entirely straightforward. Certainly a quantity of copper can survive loss of what might be regarded as the characteristic sensible features of copper, since it exists under many disguises in different combinations. But can a quantity of stuff change its natural kind? Can a quantity of copper cease to be copper without ceasing to exist? If it can, then the principle that there is a tight link between individual and species does not extend to the chemical kinds. Given the explanation of this principle in the biological sphere, that would not be particularly surprising.

As Boyle reasonably speculated, compound stuffs, like machines, can come apart.

Since a logical property of quantities of stuff is that they can survive dispersal, there would seem reason to allow that, unlike unitary objects, they can survive decomposition, e.g. that a quantity of water can survive the separation of the hydrogen and the oxygen which make it up. This conclusion may seem paradoxical, for what is the destruction of water if not its decomposition into elements? When hydrogen and oxygen are appropriately combined, it seems correct to say that a quantity of water is *made*. Unfortunately correct speech is not much help here. We are prepared to say that ice is *made* when water is frozen, yet ice just *is* frozen water, as a musician just is a musical man. We can say too that ice has ceased to exist when it has melted, yet that which was the ice has not ceased to exist. Just as it would be wrong to distinguish ontologically between the frozen-water or ice on the one hand, and the water which is frozen on the other, so it seems wrong to distinguish between the water on the one hand, and the hydrogen and oxygen on the other. Of course, the former (in each case) *is composed of* the latter, and sometimes this relation does not involve identity, as when a ring is composed of some gold. But when one quantity of stuff is composed of another, the two are surely identical. The water at any time is identical with the hydrogen and oxygen in combination, and the latter is identical with the separated hydrogen and oxygen at some later time. Talk of the water's ceasing to exist is a *façon de parler*. To take it seriously would be to accept an ontology packed with temporarily overlapping quantities of stuff. It seems preferable to think in terms of something which survives, as a quantity, through change, and leave it to physics to fill in the details: 'that, which was Grass to Day, is to Morrow the Flesh of a Sheep'.¹⁴⁷ Locke thought that the 'something' is corpuscularian matter, or something like it.

That leaves an interesting question. What is the ground of the difference between 'ice', a compound term, and 'water', a sortal? It may not be as great, logically speaking, as the difference between 'musician' and 'man'. There certainly is a difference, and the difference is grounded in reality. It is something to be found out, or confirmed, in the nature of freezing and of chemical composition. It is not just a matter of language that 'water' is the more profound classification, and that ice is not a kind of water, but water in a certain state. Yet in the case of stuffs there seems nothing quite comparable to the argument from identity to enable us to distinguish between compound terms and sortals. 'Cripple' is indubitably compound, because a cripple may well have existed before being a cripple, may possibly continue to exist after ceasing to be a cripple, and might in more fortunate circumstances have existed without ever having become a cripple. If 'cripple' cannot here be replaced by 'man' *salva veritate*, that is a sign that *man* is a sortal concept. A comparable criterion rigorously applied to stuffs would seem to prevent even *water* from being sortal, pointing to the conclusion that all stuffs are *syntheta* except whatever is ultimately conserved in physical change (call it 'matter'—but if what survives is 'energy', which is hardly a stuff, there is a problem). That rather odd conclusion may seem consonant with the doctrine of the *Essay*, but the route by which it is reached is not. For one thing, Locke held that even the idea of matter is a complex idea for one who does not know its essence. More importantly, he denied any significant distinction at all, not only between *water* and *ice*, but also between concepts like *man* or *horse*, and concepts like *cripple* or *palfrey*. Yet there evidently is such a distinction, and

one crying out for explanation. The present point is simply that the application of such a distinction to stuffs is problematic. More work needs to be done.

Locke on the difference between substances and modes

Locke's account of our ideas and names of the sorts of substances was more than a theory of the classification of substances. It also served as a part of his explanation of the difference between substances and modes, and so of the difference between first-order noun-predicates and broadly adjectival (including verbal) predicates. In the latter role it seems quite unlike the modern theory of 'natural kinds'. For although the modern theory can be very well illustrated through a discussion of substance terms like 'gold' and 'horse', it is just as applicable on the other side of the line dividing substances from modes, where it can draw on examples like 'malaria' or 'combustion'. Such a difference of scope calls for explanation, since Locke's theory and the modern theory seem to bear a real relation to each other in that, despite their differences, each assigns a special role both to exemplars (or, as Locke called them, archetypes) and to the presumption of a deep connection or resemblance between members of a species.

The classification of diseases (a subject rather surprisingly not directly discussed in the *Essay*, given Locke's medical interests)¹⁴⁸ supplies one of the more popular illustrations from which the modern theory of natural kinds sets out. Diseases are certainly 'modes', but the name of a disease will normally be introduced, and then be generally applied, on the basis of repeated experience of a set of symptoms, and on the assumption that on each occurrence they have the same common cause, whether a microbe or an underlying physiological condition. That assumption may come to be thought false, as has perhaps happened in the case of 'rheumatism'. As in biological classification, the name will then normally drop out, or be employed for a newly identified *classis*. If, however, the name is (like 'rheumatism') retained, but linked rigidly to the old set of symptoms, it becomes a descriptive or practical term like 'headache', which is not the name of a 'natural kind' in the sense of the modern theory. In another kind of case, it may be that a disease occurs, but without the characteristic symptoms; just as a substance might turn out to be a metal, since it has the fundamental nature of metal, even though, for some special reason, it lacks many of the properties characteristic of metals.

Locke's explanation of the substance-mode distinction is perhaps most plausible as long as we keep to examples consonant with his concern with the *a priori* or abstract sciences of ethics and mathematics. Like the mathematical interest in triangles, the human interest reflected in the concept of *glory* or *ambition*, of *property* or a *lie*, is not that interest in the nature of things as they are in themselves which is the interest of the natural scientist.¹⁴⁹ The customs, conventions and purposes of a society, as Locke noted, determine how its language slices up the realm of human life and behaviour.¹⁵⁰ With

respect to *natural* occurrences or processes, however, Locke himself (as we have seen) stressed the contrast between our idea of a process, which is restricted to what is overt, and the unknown 'action' or *modus operandi*.¹⁵¹ Yet it is just this underlying 'manner of operation' which can perform a semantic role similar to the 'real essence' or deep connection between members of a kind of substance. It is difficult to see how there could be a science of such modes as *pleurisy* or *freezing* which was not natural science. Did Locke then simply take a track which leads away from the problem of describing the peculiarity of names of *substances*, towards a response to some other, more general problem of *natural kinds* spanning both substances and modes?

The thought that the problem of natural kinds cuts right across the category distinction between substances and modes may be reinforced when we consider a host of first-order noun-predicates, designating things and stuffs, which are no part of a system of natural classification. Locke explicitly noticed the noun form of one class of such predicates, those 'names of relations' which include 'professor', 'citizen', 'inhabitant' and so forth, but he claimed that it is no objection to his scheme.¹⁵² Another class, that of broadly functional nouns such as names of artifacts, is significantly less easy to place outside the class of 'names of substances'. On the one hand, such terms might seem relational, since what makes something a *table* or a *knife* is a certain relation to the people who use it and designed it for use. What makes something a *statue of Churchill* is its relation to the intentions and actions of the sculptor and, through them, to Churchill. Yet on the other hand the very being, unity and identity of a table or statue may seem (and has traditionally seemed) to be bound up in its being a table or statue. 'A table', 'a knife', 'a statue of Churchill', may well appear the most explanatory answers in each case to the question, 'What is this thing?' Whereas straightforwardly relational nouns are easily enough treated as compound terms ('father' means *animal which is male and has offspring*), 'table' may seem different. No other sortal is embodied in *table* as *animal* is embodied in *father*, offering a more fundamental answer to the question 'What is it?' The term 'person' raises analogous issues. It does not designate a natural kind, if only because there is no reason in principle why beings unrelated to men or even to animals on earth should not possess the intelligence to count as persons. But for reasons a bit like those pertaining to 'table', 'person' may still at least *seem* to be a substance-term rather than a compound term.¹⁵³

According to this whole argument, the names of substances and modes might seem capable of regimentation along the lines of Table 1.

It would then appear that Locke simply picked on opposite corners of the square I-IV, contrasting natural substances with non-natural modes and taking this contrast to illustrate the general difference between substances and modes. In that case he treated problems which cut across each other as if they were the same problem.¹⁵⁴ Worse still, with all his talk of nominal and real essence it would seem that he failed to say anything of significance about the substance-mode distinction. In this section and the next it will be argued that such a proposal, initially plausible as it may be, is both unjust to Locke's thinking and inadequate to the philosophical issues.

Table 1 Names of substances and modes

	<i>Natural</i>	<i>Non-Natural</i>	<i>(Compounds)</i>
<i>Substances</i> }	I horse	II table	V father
	man	statue	boy
	gold	person	albino
	water	fuel	ice
<i>Modes</i> }	III freezing	IV truthfulness	
	measles	lie	
	photosynthesis	democracy	
	schizophrenia	agriculture	
		property	
	triangle		

First, is it true that Locke simply ignored or overlooked the contents of categories II, III and V of our table? To start with category II, it must be admitted that he offered no full-scale, set-piece treatment of artifacts, but he did say a number of philosophically motivated things about them. In a passage already discussed he explicitly assimilated the sorts or kinds of artifacts to natural species, a deliberate move which is unsurprising enough in a philosopher who wished to present natural objects as natural machines.¹⁵⁵ The main difference is here supposed to be that the artificer typically knows by what mechanical means his purpose is achieved, and so can define the terms ‘watch’, ‘clock’, etc. by some sort of real essences, i.e. by the structural attributes lying behind the powers of telling the time, striking the hour and so forth which must constitute nominal essences for the rest of us. The argument tends, of course, to play down that difference from natural species which lies in the point that the classification *watch*, just because it is answerable to a human purpose, is not, like a natural classification, liable to the requirement that there should be a ‘deep’ natural relation or resemblance which is independent of human purposes. That was not mere oversight, however. For one of Locke’s points was that natural classification and the classification of artifacts are alike just because in both we are prepared arbitrarily to ignore the overt evidence of differences in the ‘wheels and springs within’. The same comparison occurred in Boyle, who assimilated the rough and ready principles of existing natural classification to the rule that, ‘whether a bullet be silver, or brass, or lead, or cork, if it swing at the end of a string, ’tis enough to make it a pendulum’.¹⁵⁶

Mechanists also explicitly played down the distinction between natural and artificial origins, a distinction of the last importance for the Aristotelian doctrine of the form as efficient cause. According to the latter doctrine, the form exists before the artifact itself, as the content of the artificer’s purpose. The power of cutting in a certain way exists intentionally in the mind of the axe-maker, who then imposes this form on matter by bringing into existence the properties which it entails as conditions. The axe is seen as a sort of mind-dependent entity or unity, so that no independent substantial form or real

universal need be postulated such as is required to explain the existence of a naturally unitary thing. For the mechanist, however, there is no ontological distinction to be made between the natural and the artificial just because the physical nature of a thing is independent of its origins.

Nevertheless Locke elsewhere made another interesting suggestion as to the mind-dependence of the identity of artifacts, a proposal which will be discussed in Part III of the present volume.¹⁵⁷ Nor is that the only passage in the *Essay* expressing, in effect, some kind of recognition of the differences between categories I and II of our table. The doctrine that 'person' is not a name employed in natural classification, but a 'forensic' or ethical term, played a significant role in Locke's treatment of personal identity, even leading some of his followers to claim that for Locke a person is really a mode. That point too will be pursued in Part III.

The class of compound terms, category V of our table, is again one which Locke can hardly be accused of having merely overlooked, since we find so many echoes of the Aristotelian conception of composition in his own theories. Many compound terms, however, not only *father* but even *boy* (and perhaps *musician* or *baker*) could be dealt with, on Locke's principles, as relations. In fact he explicitly denied that such terms are in this sense compound, i.e. that *father* embodies *man*. His reasons were characteristic and interesting. Although he agreed that the idea of a father is 'a notion superinduced to the Substance, or Man', he did not draw the conclusion that it *includes* the idea of man. It 'refers only to an act of that thing called Man; whereby he contributed to the Generation of one of his own kind, let Man be what it will'. It applies extrinsically to men, but not *qua*, members of that particular species. Hence those 'who have far different *Ideas* of a *Man*, may yet agree in the notion of a *Father*', which would be impossible if *father* included *man*.¹⁵⁸ Locke was interested in the possibility of clear and distinct ideas of relations as objects of demonstration and 'science', and he needed it to be possible for someone to give a real definition of *father* without being able to give a real definition of *man*. The argument collapses, of course, on a different view from his as to how words like 'man' possess meaning. Yet Locke's conclusion that the real object of the concept *father* is an abstract relation, not the substantial bearer of the predicate, the thing related, is not entirely implausible. The study of *policemen* can be thought of as the study of a role, rather than the study of a sub-class of the species *man*, or the study of *man* in a certain state. Nor is it implausible that the 'science' in which the concept of this role (or even the concept of the status, *fatherhood*) might figure is *a priori* moral and political theory, a science of rights and duties, rather than sociology conceived of in the modern way as a branch of quasi-natural science. Other examples, no doubt, are less favourable to this model. The object of study or inquiry about *adolescents* or *the adolescent* is surely the species man at a certain stage of development (perhaps, in our society): not, as Locke's account of 'young' and 'old' would suggest, mere relative duration.¹⁵⁹

Not all 'compound terms' can be represented as relational, however, and the most telling criticism of Locke in this respect must lie in his lack of resources to distinguish predicates like 'man' and 'water' from predicates like 'ice', 'albino' or even 'cripple'. All are for him complex, and rigidly linked to criteria of application. Now as for the logical difference between 'water' and 'ice' (which Locke expressly denied),¹⁶⁰ the Kripkean

account of the names of 'natural kinds' seems to have little to say about it either. If 'water' names a natural kind of substance, and 'freezing' names a natural kind of process, then presumably 'ice' names a natural kind too, as far as all Kripke's criteria go. Yet ice is not a species or even a 'kind' of water. Much the same point can be made about 'albino', although perhaps not about 'cripple', since there are many causally dissimilar ways of becoming crippled, not to speak of different ways of being so. It seems, then, not only that the modern theory will not by itself explain compound terms, but that the problem has simply not impinged on the thinking of proponents of the modern theory as it had on Locke's thinking. For the denial of a logical difference between 'man' on the one hand and 'albino' and 'cripple' on the other is a nettle which, as we have perhaps sufficiently seen in previous chapters, Locke grasped defiantly and knowingly as a part of his campaign against Aristotelian species. For example, he argued that 'a man is a rational animal' is capable of being necessary only in the same way as 'a palfrey is an ambling horse' is necessary (i.e. by nominal definition), and this argument openly and, surely, deliberately implies that 'palfrey' is as much a name of a species as is 'horse'.¹⁶¹ Any predicate, he asserted, provided that it is annexed to a nominal essence, is potentially as good an answer as any other to Aristotle's question, 'What is it?'¹⁶² Locke was here certainly wrong, but he was not guilty of mere oversight. He had a theoretically motivated position on the matter.

To return to the different kinds of modes, the topic is in fact rather more complex than the simple opposition between categories III and IV of our table would suggest, especially if we bring Lockean simple ideas into consideration. A very large class of non-substances, including what may be called phenomenal modes (as redness and pain), psychological modes (as love and memory), and functional or dispositional modes (as elasticity and adhesion) seem to fall somewhere in between the two categories. The presumption of a common underlying nature does not play the same semantic role for such terms as it plays for items in category III, but the existence of a deep resemblance is not as incidental as it would be for items in category IV. We do normally and naturally suppose that there is a deep causal resemblance between cases of modes falling into this intermediate class. It would be something of a surprise to learn, for example, that two or three very different mechanisms share the responsibility for phenomenally indistinguishable human pains, or for memories of past experience; or that several physically quite distinct kinds of light are seen as the same shade of brown. Yet such discoveries (which do sometimes occur and could in principle occur more generally) do not undermine the use of the concepts in question. We might be led by them to talk of two *kinds* of pain, or of memory, or of brown, but without the imputation that there is something inadequate in the concept or ambiguous in the name. The principle of their generic classification as 'pains', 'remembering' or 'brown' would remain untouched, being independent of the principle of their specification on the basis of different underlying mechanisms. At the same time (as the long discussion, in Volume I of the present work, of the notions of *perception* and *knowledge* sufficiently illustrates) it is a satisfaction-condition even of generic terms of this type that *some* specific mechanism, known or unknown, be involved. 'Memory' without a specific mechanism would be accidental, and so not be memory, although specific mechanisms in the same or different

animals might be as different from one another as chalk from cheese, or chemistry from electronics. Much the same goes for (say) stickiness or elasticity. There must be some specific mechanism to link the appropriate observable changes in an elastic object, and changes in its relations to other objects, with an underlying state of the object. But there is no reason *a priori* why the mechanism should be the same in all elastic objects. In this case, indeed, it might be rather surprising if it were. The case is different, however, with items clearly in category IV. We might perhaps speculate, or have good reason to expect, that there will be a deep explanation in common to instances of *violent crime*, or *democracy*, or even *festivals*, or at any rate that the specific explanations will be relatively few. Yet it is not a satisfaction-condition of a term like these that any particular one of a number of specific mechanisms should be involved. Lying is lying, as triangularity is triangularity, whether or not any pattern at all emerges from a study of its causes; or, if a certain pattern does emerge, whether or not the particular case in question falls within that pattern. The satisfaction-conditions of such terms lie entirely elsewhere, in spatial relations, in intentional content, in pure performance and the like.

Locke's approach to these distinctions among non-substances (which I have called 'modes' quite generally, extending his own use of that term) was not as straightforward as it might be thought from some of his better-known remarks about mixed modes and relations. For one thing, some of the examples of natural modes which I have given would obviously fall into his class of 'simple ideas', a class which includes at least some powers. In their case an effect is supposed to be employed to signify an unknown cause, and the presumption of such an underlying common character is taken to be intrinsic to their names' possessing significance. But setting such cases aside, Locke's argument can seem to us exceedingly unclear. On the one hand he was prepared to argue that our interest in complex human actions and powers is precisely *not* in underlying resemblances, but rather in ethics, in action as 'the great business of Mankind, and the whole matter about which all Laws are conversant'.¹⁶³ On the other hand, even while stressing the supposed difference from substances, he was apparently prepared to allow that a presumption of an unknown *modus operandi* is after all at least sometimes involved in the formation of such an idea. Once more on the subject of the facts of life, he wrote:

For if I believed, that *Sempronia* digged *Titus* out of the Parsley-Bed, (as they use to tell Children,) and thereby became his Mother; and that afterwards in the same manner, she digged *Caius* out of the Parsley-Bed, I had as clear a Notion of the Relation of Brothers between them, as if I had all the Skill of a Midwife; the Notion that the same Woman contributed, as Mother, equally to their Births, (though I were ignorant or mistaken in the manner of it,) being that on which I grounded the Relation; and that they agreed in that Circumstance of Birth, let it be what it will.¹⁶⁴

Now a true explanation of the possibility that someone might have such ridiculous beliefs about brotherhood while still possessing the 'idea' of *brother* would doubtless have to appeal to the conception to which Locke was so opposed, that of a 'division of linguistic labour'. The midwife *is* in a special position.¹⁶⁵ But the main point relevant to our

present argument is that, even in trying to support the notion that a concern for relations is a concern solely for what is superficial and knowable, he slipped into a model hardly different from his account of ideas of substances. And if it is true, as he insisted, that we can imagine the idea of *adultery* being invented before it was instantiated, we surely cannot imagine such an origin for the idea of *birth* or *brotherhood*. This is especially true if such an idea were inexplicit about the mechanisms of generation, for then especially (as even with an idea like *adultery*) it would be necessary to gesture deictically, if indirectly, towards the real, but unknown mechanisms. Otherwise, almost *any* relation could be meant.

If that was all Locke had to say about ideas of modes, criticism of his theory would have to stop here, with his being convicted of the mistake of overlooking ‘natural’ modes or denying them significance. Yet he had a much broader conception of the status of modes than appears if we focus exclusively on the doctrine that, *because* ideas of modes and relations are in some way abstract or constructed or mind-dependent as substances are not, mixed modes and relations are *therefore* proper objects of *a priori* or abstract science rather than of natural philosophy. The conclusion, as we have seen, is not sound, just because there are ‘natural’ modes and relations. It is easy to think that this means that the premise is not sound either. Yet the false doctrine that ideas of modes and relations may in general be formed without reference to reality was only a part of Locke’s argument for that premise. More fundamental, it seems, was his view that to assert the existence of a substance is to claim that its defining attributes coexist ‘united in one subject’, i.e. enjoying some kind of natural unity not enjoyed by modes. Someone’s jealousy is not unitary in the way that the man or woman who is jealous is unitary. Locke’s claim was that the formation of ideas of mixed modes is peculiarly arbitrary and undictated by nature even when the combination of simple ideas in question does occur in our experience. I believe that that claim can properly be extended to what I have called ‘natural’ modes. One reason why it is difficult to recognize such a possibility from Locke’s arguments alone is that in the *Essay* the claim is almost universally illustrated by examples of non-natural modes. Indeed, it is fairly evidently conflated in Locke’s mind with claims which *depend* upon such examples.

Thus, interwoven with the proposition that in forming ideas of mixed modes the mind does not tie ‘itself to a precise imitation of any thing that really exists’, we find the argument that, even when an idea of a mixed mode *is* formed from nature, there are indefinitely many other, equally deserving concomitances which we feel no compulsion to unite in a complex idea with a name.¹⁶⁶ This point could well be illustrated by an example of a natural mode: we have the concept of a *storm*, but no special concept for (say) the change of cloud-formations from cumulus to cirrus. A typical illustration of Locke’s, however, is that the killing of a man merits the name ‘murder’ and the killing of a father, ‘parricide’ (i.e. ‘patricide’), but there are no special names for the killing of a sheep or of a son. Similarly, when he wanted to say that by bestowing a name on a mode we bestow a greater reputation for unity than nature can possibly justify, then almost without exception his examples were, quite unnecessarily, non-natural modes:

Thus the Name of *Procession*, what a great mixture of independent *Ideas* of

Persons, Habits, Tapers, Orders, Motions, Sounds, does it contain in that complex one, which the Mind of Man has arbitrarily put together, to express by that one Name? ¹⁶⁷

It can help us to see why such one-sided illustrations were attractive to Locke if we notice another element in his argument. The arbitrariness of ideas of mixed modes, he claimed, is not 'without Reason', for their construction is 'done by the free choice of the Mind, pursuing its own ends'. ¹⁶⁸ Our choice always reflects some human interest or practical purpose, together with the general need to communicate. Since such interests are largely determined by a society and common to its members, people tend only to acquire ideas of mixed modes corresponding to the words already existing in the language of their society, and seldom feel the need for more. Now it seems likely that Locke associated the point that certain concepts exist for practical and social reasons with the thought that they are practical and social concepts, especially since the latter tend neatly to illustrate the thesis that certain concepts are practically and socially determined. There exists a bad argument in modern philosophy against the fact-value distinction which proposes that even concepts of natural science are 'evaluative' or 'value-laden', or 'express values', since even scientists possess the concepts they do for practical and institutional ('ideological') reasons. Something more sophisticated sometimes lies behind this argument, but as it stands it is a blatant *non sequitur*. If the explanation of our having the concept of a *storm*, but not some other possible meteorological concept, lies in our practical interests, that does nothing to show that the concept *storm* expresses or concerns practical human interests or values in remotely the same way as do the concepts *glory* and *ambition*. Perhaps something a little like this confusion exists in Locke's argument, sophisticated as his general employment of the substance-mode distinction may have been.

A rather more profound source of Locke's theoretical neglect of natural modes is discernible in his conception of the unity of a substance. The contrast always at least implicit in his remarks about the arbitrariness of our ideas of modes is with confrontation by some hitherto unknown substance which, independently of our interests or the already existing language, *demand*s the introduction of a name. This contrast was not easy for Locke to express since he was also arguing (not without reason) that there is much that is arbitrary in the singling out and naming of the sorts of substances. Yet that in a way only underlines the *greater* extent to which in his view the singling out of modes must be the work of the mind. In the case of substances he held that we arbitrarily set the boundary of the species by selecting a nominal essence from an indefinite number of coexisting qualities and powers. Yet our choice is considerably circumscribed in that we choose only from those attributes which are presented in our experience as the attributes of one thing, i.e. as emanating from a single source, the 'cause of their union'. This was Locke's explanation of the unity which, as he plausibly argued, a procession, a triumph and an apotheosis obviously lack. But, if correct, the point seems capable of being drawn from natural modes as well as from non-natural ones. A storm, an earthquake, an epidemic, do not seem each to be any more 'one thing' than a procession, at least not in the way in which a horse or a piece of gold is one thing. Now it is because modes in general lack the 'given' unity of substances that in the formation of ideas of modes our choice was

supposed by Locke to be entirely unconstrained. Yet, as we have sufficiently seen, he tried to explain the unity of substances by the role of the underlying real essence. For qualities and powers to coexist in one subject just is, on his view, their flowing from a general substance (such as matter) constituted in a certain determinate way, i.e. their deriving from a unitary ‘real constitution’ or essence. So if modes lack the real or given *unity* of substances, it might seem that they must *ipso facto* lack underlying or substance-like real essences too. And if, for modes, no unity is given naturally, then all unity must be imposed arbitrarily, at the level of the nominal essence: ‘in mixed Modes, the unity necessary to any Essence, depends on the Mind’.¹⁶⁹ Consequently Locke’s doctrine that mixed modes have, of their very nature, no real essences behind their nominal essences is just what we might expect from his theory of substance and substantial unity. Moreover, if the real essence of a specific substance is the determinate constitution of the matter which comprises it, it may seem inevitable that modes, which are not composed of matter, should lack underlying real essences.

Locke’s accounts of substance and of the names of substances have been criticized above, and there is no need to agree with his accounts of mixed modes or of unity either. But if, under the influence of the arguments of Kripke and Putnam and in the face of examples like *freezing* and *measles*, we wish to recognize a class of natural modes with names semantically comparable to the names of natural substances, we still need to form some assessment of Locke’s claim that modes differ from substances in respect of unity. If we interpret that claim in terms of his theory of substance and real essence, then we must presumably reject it. In having an underlying ‘real essence’ a natural mode such as a storm or an attack of chicken-pox will presumably possess as such all the natural unity that a real essence can bestow. On the other hand, there must be *some* intelligible difference between substances and modes, and the intuitive claim that a horse is a naturally unitary, discrete thing as an earthquake is not seems as good a point as any to start an investigation of that difference. At any rate, Locke’s traditional proposal that natural unity and distinctness is the prerogative of substances is one that should not be allowed simply to fail by default, or to run into the sands of modern conceptualist prejudice.

Although an attack of chicken-pox may seem substance-like not only in possessing something like a Lockean real essence, but also in having a history for the most part internally determined by its earlier stages, yet there is no denying that its individuality is imposed by us, its boundary-principles or existence-conditions are set by our concept. To see that this is so, consider a variety of possible candidates for the status of being a case of chicken-pox. In different subjects, let us suppose, the virus is present as follows:

- (1) without symptoms, beyond positive results to certain tests
- (2) with minor characteristic symptoms (such as swollen glands and a slight ‘cold’) but without ‘illness’
- (3) with fever, and so ‘illness’, but with no other characteristic symptoms
- (4) with a spot or two, but no fever
- (5) with fever, spots and other characteristic symptoms (the paradigm case)

In which of these cases does the subject have chicken-pox? Perhaps we exclude (1) and

(2), since they are not cases of illness. We let in (4), however, since the spots are so characteristic. That does not tie 'having chicken-pox' to symptoms, without regard to a presumed common aetiology. It is just that the 'real essence' here, a certain virus's acting to a certain effect, owes its boundary-principles in part to the general, functional notion of being ill, if not to some more specific determination of the effects.

Here we may consider how pragmatically the genus of 'diseases' is determined. Yet this general notion essentially helps to determine, not only the boundary-principles of the species, but also the boundary-principles of the individual. In this respect the genus 'disease' is quite different from the genus 'animal', for the latter is part of a hierarchical system logically (even if not, in some minor respects, genetically) built up from below, *from* the level of particulars. A mark of this difference is that problems about the boundaries of the genera 'animal', 'mammal' and so forth are in general quite distinct from problems about the boundaries of the species. If there is a group of individuals which, after exhaustive exploration and appraisal of their characteristics, continue to give rise to doubt as to their membership of the species *mountain gorilla*, their problematic status will not be due to doubt as to whether they are animals, or even whether they are primates. In general, anything which is, from the viewpoint of taxonomy,¹⁷⁰ possibly a mountain gorilla will certainly be a primate, except in so far as it may be questionable whether mountain gorillas are primates. But it may be questionable whether someone has chicken-pox just because they are not clearly classifiable as ill. It may be beyond question that, *if* they count as being ill, then the attack of illness is an attack of chicken-pox.

The imposed or conceptual character of the boundary of an attack of the disease shows itself again if we consider, for example, the place of the individual episode in the history of the victim, its relation to other aspects of the victim's general condition and so forth. If the virus responsible for some disease is commonly present in the body before and after an attack, and is liable to cause a recurrence, then the attack appears as a merely striking slice, which is bounded only in so far as it is striking in a way that matters to us, out of a continuing process in the victim. If the illness occurred in the context of a general breakdown of the victim's health, it can equally appear as a part or aspect of that other process. If the victim infects others, the process taken to encompass what goes on in the individual may have the wider dimensions of an epidemic, or of the history of a strain of virus. Every episode is a part of indefinitely many wider episodes, and there are no natural wholes or parts, no 'given' articulation of episodes. It is our concept which determines that the smoke is an effect of the fire, while the flame is a part of it. The only reason why the breeze now blowing over England is a result, rather than an element of the hurricane out to sea may be that it is too gentle to count for us as the latter. By contrast, a horse is by nature a whole, its undetached hoof is by nature a part of that whole.

The *Essay* contains another traditional ground of distinction between substances and modes relating to unity. Many twentieth-century logicians and philosophers of science have been prepared to argue that material objects are in effect long events, processes which take time to complete. Locke, however, seems with reason to have regarded it as peculiarly evident that processes are mind-dependent. They are 'fleeting and transient combinations of simple *Ideas*, which have but a short existence anywhere but in the

Minds of Men'.¹⁷¹ His examples are here a *triumph* and an *apotheosis*, and he meant, not of course that they are momentary, but that, 'being Actions that required time to their performance', only a momentary, 'fleeting' *part* of them could exist at any one time. They 'could neither of them exist altogether any where in the things themselves', and it is the *name* which unites these fleeting parts over time. Of homoeomerous or stuff-like 'actions', 'v.g. Motion and Thought', he elsewhere noted that they 'consist in a continued train of succession, each perishing the moment it begins'. Since there is a succession of 'parts', 'each having a different beginning of Existence', there is no question of the action at one time being identical with the action at another.¹⁷² The immediate point here seems to be that my walking at t is never identical with my walking at $t+1$. Triumphs and apotheoses, however, are like walks rather than walking (if less homogeneous than either), and my *walk* at t may be identical with my *walk* at $t+1$. Here Locke's express doctrine is that we *impose* identity over time, uniting the successive parts of the mode by a single idea with a name.

The contrast implicit in these remarks of Locke's is, of course, with substances, which are *not* composed of fleeting parts. The substance does 'all exist together', and its continuity is real rather than imposed or constructed. The immediate logical point can be illustrated, in modern terms, with reference to so-called 'phase-sortals'. The child is not, as modern logicians have been inclined to allege, a temporal stage or phase or part of the human being, but is the whole human being: indeed, the adult itself during an early stage of its life. The performance of the first movement, however, is only the first part of the performance of the symphony. If, after five bars, fire breaks out and the orchestra stops playing, there has been no performance of the symphony but only of the first five bars. A performance of the symphony *would* have taken place, if fire had not broken out. By contrast, if a child dies soon after birth, the individual man has already existed. It is not that a different individual, an adult, *would* have existed later, in happier circumstances. What has been curtailed is a life-history, a peculiarly parasitic mode, not the substance itself.

The significance of this contrast is worth exploring both for its own sake and because it might help to explain why Locke was not impressed by the analogy between natural modes and relations on the one hand and natural substances on the other. First, the 'real essence' of a substance is the physical constitution from which flow, not merely actual effects, but the indefinite multiplicity of coexisting powers which, Locke argued, stands in the way of our having either an adequate or an agreed idea of the substance: 'not having tried all the Operations of all other Substances upon it, and found all the Alterations it would receive from, or cause in other Substances'.¹⁷³ By comparison, our idea of a mode requires 'nothing else to make it perfect'.¹⁷⁴ It may of course be objected against Locke's formulation of the contrast that ideas of modes are not necessarily in his sense 'perfect'. We may carefully observe and identify more symptoms of the same disease, and different doctors may employ better or worse diagnostic criteria. Yet it is the conception of an object open to unlimited testing which is in question. What prevented Locke from giving due weight to the analogy between the *modus operandi* of a causal process and the real essence of a substance was doubtless his desire to give a general explanation of the *a priori* sciences, together with his corresponding predilection for the

comparison (both analogy and contrast) between the powers flowing from the unknown real essence of a substance and the geometrical 'properties' flowing from the known definition of a figure. Yet it seems true that the model of an enduring thing with an unknown structure, capable upon test of affecting other things in an indefinite variety of ways and of suffering an indefinite variety of changes itself, has no clear application to modes or relations however 'natural' they may be.

The attribution of powers and causal properties to a substance is (as Locke everywhere supposed) close in its character to hypothetical assertion. We can correspondingly associate the reluctance of particular modes to support powers with the fragility of their identity under the stress of hypothesis. Selecting from the past an individual man, for example Napoleon, we can intelligibly say that, if circumstances had been different at some period or moment of his childhood, he would have lived out his life in Corsica as a deformed idiot, or that he would have died a boy. The supposition does not in any way call into question the identity of the man in the 'possible world' of which we speak, or frustrate our intention to refer to Napoleon. There are no constraints except the limits of physical possibility upon our supposing that an individual substance, a thing or a particular quantity of stuff, could have enjoyed a different history and yet have remained the same individual. Yet if we claim that a certain action by Napoleon at an early stage of the Battle of Waterloo would have led to an immediate retreat of the allies, or to a battle with a different course and outcome, then it seems that we have in each case imagined a possible world in which the individual event, the Battle of Waterloo, simply (or at least arguably) did not occur at all. In both suppositions only a part of the actual battle, the opening skirmish, occurred. In the second, a numerically *different* 'Battle of Waterloo' occurred.

In other words, the logic which seems to make clearest sense of such suppositions about events or processes involves the divisibility of processes into temporal parts, taking the supposition to draw a line through the whole process at a point in time. The earlier part exists, while the later part does not exist, in the supposed possible world. The principle that two events are identical if and only if they have all parts in common thus holds between possible worlds, as we might expect from the impossibility that temporal parts should be replaced. It may seem extravagant to maintain this principle to the extent of claiming that, if we suppose the slightest divergence from the actual course of an event, then we suppose a different event. Yet it is difficult to see any logically pure ground for stopping short of that claim. But even if we are seldom so strictly logical in ordinary speech, the point of contrast with substances is clear. In the case of substances, as it is widely recognized, there are comparable restrictions on the supposition that an individual might have had different origins. We cannot intelligibly speculate, for example, that in certain circumstances Napoleon would have been the child of different parents. But there are no such restrictions relating to the *consequent* life-history of a substance. There is nothing in the least problematic in respect of identity about the claim, 'If it had been given proper conditions, that pickled egg would by now have been a fine hen.' All this is involved in the point that a substance *endures* rather than *lasts*. It is, moreover, the condition of the explanatory role of substances as the subjects of powers, a role celebrated in their different ways by both Aristotle and Locke. We can identify all at

once a potential subject of indefinitely various and perhaps, for us, unpredictable change: a subject to be tested and tampered with, to be 'knocked and beaten with flints' (as Locke imagined a new-found substance being treated by Adam in primitive pursuit of experimental philosophy),¹⁷⁵ or to have its hitherto unknown 'normal' life-history, with all its metamorphoses, however exotic and unexpected, observed and recorded.

It is interesting to compare with this common-sensical model the monadism of Leibniz, who combined in his idiosyncratic conception of a substance characteristics of both substance and mode. On the one hand he eviscerated the capacity of substances to endure or exist through change by disallowing accidental changes. Like occurrences, monads have, as it were, a course every twist and turn of which is essential to the individual in question. The paradoxical consequence is that the world is presented as a static system in which nothing *happens* to the individual substance. The latter appears as a four-dimensional sausage stretched out in time, individuated and bounded by the concept of the unique life-history which it necessarily fulfils. From Leibniz's point of view, on the other hand, that is not because monads are event-like, but because they alone are truly substantial. The Leibnizian substance exists all at once in a very special sense, since *all* its attributes flow from its being, so that it is from the start 'pregnant' with its entire future. Its life-history is the unfolding of the only possibility for this thing, and the monad is so discrete from other things that it is never in the least affected by them. Yet it may be objected on behalf of a more orthodox view that we understand how the future of an individual is 'contained' in its present only in terms of the embodiment of potentiality in what is actual and relatively stable. And we understand embodied potentiality as something the actualization of which depends on the circumstances as well as on its physical embodiment. In other words, it is essential to the category of substance that individual substantial things (whatever their inherent tendencies) are capable of any of an indefinite number of life-histories in different possible worlds. To assimilate substantial things to occurrences with temporal parts is to deny the category this fundamental logical property.

Unfortunately, however, the opposition between *existing all at once* and *having temporal parts* seems inadequate to distinguish substances from all temporally lasting modes. Despite what has been said, it does seem possible to conceive of processes and activities, if not of events and occurrences, as being more substance-like than we have considered: as being, as it is sometimes put, 'continuants'. An attack of a disease, for example, can be thought of at its onset as possessing potentially a normal course, which may, however, be cut short or made less severe by treatment, or otherwise modified by circumstances, without affecting the identity of the attack. We may be drawn by the standard logic of occurrences to distinguish between the actually mild hepatitis-episode which has gone along with the administration of gamma-globulin and the severe possible episode which would have occurred without treatment. Yet something other than mere logical laxity may lead us, on the contrary, to treat the attack as if it existed, like a substance, 'all at once', as a subject of various potentialities and as a substrate of change. The significance of this conceptual possibility may perhaps be diminished by the point that the potentialities are ultimately embodied, not in the earlier stages of the disease (which is not material), but in the body of the patient and in the attacking virus. But the

neat logical distinction between events and things seems to have become blurred.

Examples which may seem even more awkwardly boundary-blurring are supplied by such entities as waves, fires and hurricanes. Each involves a sort of principle of activity which seems relatively unaffected by all but abnormal circumstances, and through which what occurs later is largely determined by what occurs earlier within the individual. Each has a normal course, which we can think of as largely 'internally' determined. As David Wiggins says of substances, such entities are not in equilibrium with their circumstances. Yet the analogy with substances is only partial, and in the case of such modes there is a conceptual choice. We can think of *a fire*, as perhaps the Fire of London is normally thought of, as an occurrence with earlier and later stages which took place and took time. But we can equally take the process of combustion itself to be the fire (the word 'process', of course, shares in the ambivalence). So understood, a fire may spread from one place to another, may be fanned fiercer by the wind, and may be wholly or partly extinguished by a bucket of water. By a third way in which we can conceive of a fire it may seem to be a true individual substance, a space-occupying, movable, enduring object, which exists all at once and is composed of the matter being consumed and continually replenished. (Heraclitus, at any rate, was impressed by the analogy.) The same three sorts of possibilities arise in respect of *hurricanes*. A hurricane may be an event with a beginning, a middle, and an end. Otherwise it may be a disturbance or activity with a spatial centre, which moves about, increases or decreases in extent and ferocity, is 'born', 'lives' and 'dies'. Finally we can conceive of a hurricane as something even more substantial, possessing a 'body' of air and cloud, with changing, replaceable material parts. There are perhaps just two ways of thinking of a *wave*: either as a disturbance in a medium or as something quasi-substantial, more or less high, and composed of rapidly changing water.

These considerations might well provoke some embarrassing reflections—embarrassing that is to say, to the thesis that the individuation of substances, by contrast with the individuation of modes, is not at all arbitrary. For it might be argued that how we think of a hurricane or a fire is, after all, a matter for arbitrary decision. It might be claimed that for each of these ambiguous predicates there are certain satisfaction-conditions common to its different senses; and that whether we think of the entity in question as an occurrence taking time, or as a mobile disturbance or activity with a life-history, or as a material thing with a body of air or gases (or whatever) united by a principle of activity, simply depends on which 'criterion of identity' (or which logic) we decide to couple with the common 'criterion of application'.¹⁷⁶ In so far as the decision seems arbitrary, it might seem that in general the 'primacy' of the category of substance for us is nothing but an ingrained habit: it is customary and for some reason convenient to associate one particular type of criterion of identity with a broad class of nouns like 'horse' or 'rabbit' which mark important principles of activity. It simply needs an effort of imagination to see that, by analogy with 'hurricanes', we might just as well think of 'horses' as occurrences with temporal parts, or as processes. Yet this argument does not explain why hurricanes and fires are rather poor individual substances as compared with horses, and the whole case loses a good deal of its seeming force if we devote more than the briefest reflection to the supposed analogy, provisionally set out in Table 2.

Table 2 Hurricane and horse

A	B
(1)The hurricane ^a , conceived of as an occurrence with temporal parts, etc.	The physical, internal life-history of the horse
(2)The hurricane ^b , i.e. the mobile disturbance conceived of as existing all at once and subject to change and to different possible futures, etc.	The life-process of the horse
(3)The hurricane ^c , conceived of as a quasi-substance composed of ever-changing material (violently moving air and cloud) united and distinguished by a principle of activity (a certain sort of swirling motion due to low atmospheric pressure).	The horse

The anti-conceptualist intuition, and the realist's dissatisfaction with fires and hurricanes as individual substances, might find expression in the claim that there is an evident disparity between columns A and B. The individuality of B3, the horse, is evidently prior to the individuality of B1 and B2. The horse is the 'given' individual, while its life and life-history are parasitically identified. In column A no individual is 'given', least of all A3. This is an appeal to the intuitive notion to which Locke also appeals in his claim that modes lack real or natural unity. The general thesis might be illustrated by the point that an event has no natural boundaries, no natural beginning or end. Some water's coming to the boil, or freezing, is just a part or aspect, sliced out by us, of the continuous physical process involved in the water's possessing variable temperature. If the nature or essence of boiling or freezing or exploding strikes us as individuating a discrete event because its coming into being involves a change of kind, not merely of degree, then that is just the way it strikes us. But a horse really is a distinct individual. It strikes us as distinct and unitary because we apprehend it as it is.

What kind of natural boundary, then, is it possible that a horse should have, but impossible that an explosion or a hurricane should have? What sort of unity and discreteness and continuity is peculiar to substances? In the next chapter an answer to these questions will be proposed, although the argument for it will be completed only with Part III, below, the critique of Locke's theory of identity. The answer is an extremely simple one, but it is not such a simple matter to show that it *is* the answer.

Reflections on the notion of substance

The notion of a category of substance, like the notion of knowledge, is both richly complex and unitary. As in the case of knowledge, ¹⁷⁷ a variety of principles can be advanced each of which reveals some aspect of what it is to be a substance. Such characterizations can seem disparate, and their very plurality commonly provokes the criticism that the word 'substance' is irremediably ambiguous. Yet reflection can reveal their interconnections, and the genuine unity of the list.

One important difference between the notions of substance and knowledge is that 'substance' is a philosophical term of art and a term, as they used to say, of 'second intention', connoting logical and ontological status. Moreover, it may seem evident that, since different philosophers have used the word in different ways, there is no correct way of using it, and so no truth about 'substance' to be uncovered by reflection on its use. Yet its admittedly rather varied employment has been part of a sustained attempt to understand (if also sometimes to improve) fundamental ways of thinking which are not only familiar to us, but which penetrate to the deep structure of language. The automatic hostility with which the traditional notion is liable to be received by present-day philosophers has disparate sources. The empiricist's horror of supposed unknowables is now, perhaps, less fashionable than the neo-idealist or conceptualist dislike of explanations of specific logical form which present it as other than arbitrary, mutable and at best useful. Another important motive is the feeling that modern physics has undermined the claims to ultimacy inherent in the substance-based ontologies of the past, just as (it is supposed) any currently preferred ontology is in turn vulnerable to future physics. The aim of the present section is to show how traditional realist doctrines of substance, suitably interpreted, have an explanatory force which pragmatism or any other form of conceptualism cannot hope to match. The explanations it supplies are of primitive structural features of our thinking to which considerations drawn from the philosophy of physics are simply irrelevant.

The present argument is closely consonant with the epistemological argument in Volume I for a distinction between the level of scientific theory and the level of pre-theoretical or experiential knowledge of the world. Roughly, the proposal is that the traditional doctrines relating to a category of substance, which were originally taken both to set out the structure of our pre-theoretical thinking and to supply an adequate schema for scientific explanation ('substances' being both the primary individual objects of experience and the ultimate objects of science), in fact has more to tell us about the former than the latter. The doctrine marks the way in which reality impinges on us at the level of experience, and so explains the contribution made by reality to primitive logical form.

Although preconceptualist realists disagree over what things are paradigm substances, as well as over the details of what it is to be a substance, it is not difficult to draw from the tradition at least a rough list of the leading characteristics of the category. Such a list of properties should even help us to map and understand the disagreements and divergences; although understanding also requires knowledge of the motives, deriving largely from science and religion, for those new emphases and reinterpretations which transformed, but in some ways weakened the doctrine of substance in the seventeenth and eighteenth centuries. Yet for the immediate purpose orthodoxy is generally more relevant than heresy, and broad principles than idiosyncratic variations. The present argument will therefore assume that such things as horses and plane trees are paradigm substances, while homoeomerous substances, such as gold and water, call for special, but integrally related treatment (treatment sketched out below under the heading 'materiality'). Attributes such as a thing's redness or squareness, and events, actions or processes such as walks, walking, thunderstorms and the like exemplify non-substances or (for the sake of a single name) 'modes'.

Most of the claims embodied in the following list have on occasion been impatiently brushed aside, or simply ignored, by modern philosophers arguing on behalf either of empiricism or of conceptualism. Some of them, however, have been reinterpreted and absorbed into this or that version of conceptualism. One aim of the present argument is to reveal the unity of the list, which will emerge in so far as each item can be explained, and needs to be explained, by reference to the others. Another aim, obviously, will be to make the principles appear tenable, although not all will in the end be recommended, or recommended as they stand.

- (1) Substances are the ultimate subjects of predication, and therefore the only beings with independent existence.
- (2) Substances are real unities (physical and logical).
- (3) Substances are material. Individual substances are distinguished from one another at any one time by their matter. ¹⁷⁸
- (4) Substances exist all at once, and exist through time, or endure. (Events, in contrast, take time or unfold.)
- (5) Substances are active, the ultimate sources of change. Their underlying natures or essences are the ultimate principles of explanation.
- (6) Only substances fall into true natural kinds, and every truly individual substance is a member of a natural kind. ¹⁷⁹

All these properties of the category of substance have been mentioned or discussed above, some at some length, and will be further elucidated below. At the same time the search for a tenable and unified understanding of the list will call for more than the elucidation of traditional views. The traditional doctrine of substance deserves to be taken more seriously, but it also needs some renovation. Let us consider the listed properties *seriatim*, both summarizing and amplifying earlier discussion.

Logical priority and independence

That substances are the ultimate subjects of predication is perhaps the most famous claim made about the category. The other categories are predicable of substance, while substance is only predicable of itself. There are, of course, particular subjects of predication not in the category of substance: the colour of this curtain, for example, which is the subject of the assertion that this brown is fading fast (universals do not fade). Yet such subjects owe their existence to the satisfaction of a predicate by something else. This brown exists just as long as the curtain is brown. It is admittedly also true that this horse exists just as long as something is a horse: yet here the 'something' is not something else, but is the horse itself. Their having no need for 'something else' as their subject is what has given rise to the ontological doctrine that substances are the only logically independent individuals, possessing independent existence.

A non-substance or mode may exist dependently on another non-substance, as beauty may belong to a shape. But then the shape, as a particular existent, will normally depend on something shaped, and ultimately every mode depends on a substance or substances. A shape, it is true, may sometimes exist as a particular without belonging to a particular thing or substance. The triangle formed by three pencils on a desk exists, although no substance on the desk is triangular. But then it exists just as long as the pencils are in a certain relation.

Logical and metaphysical dependence is taken to be reflected, if imperfectly, by the grammatical phenomenon of nominalization. Broadly speaking, general nouns which are not substance-nouns are nominalizations of predicates which are not nouns. The imperfection of this rule of thumb is easy to illustrate: 'courageous' derives from 'courage'. Other counter-examples seem more interesting, because less accidental. Although the noun 'square' may be presumed to derive from the adjective (and so, perhaps, with the French '*rond*'), 'triangular' is a grammatical offshoot from 'triangle', 'circular' from 'circle' (or '*circulus*'). Yet, even if grammar and etymology reflect logic imperfectly, the logical point is itself unaffected. Whether or not the verb 'battle' preceded the noun, a particular battle exists just in so far as certain people are fighting.

The question between realism and conceptualism is whether that logical point reveals something about the different ontological status of people and battles, or else merely reflects an accidental feature of 'our conceptual scheme'. It is true that the advocate of conceptual relativism is likely to reject this characterization of the issue on the grounds that ontology itself is always relative to a conceptual scheme: ontological assertions, it is commonly held, express preference for one scheme over others or (if talk of alternative schemes is thought to be beyond the bounds of sense) purport to reveal something about the only scheme we have. But the principles of the pre-Kantian substance-theories, including the notions of dependent and independent existence, can only be understood in terms of a conception of ontology as absolute. For the conceptualist, all beings are concept-relative: for the realist, some beings are concept-relative but others, substances, are not.

Acceptance of the famous logical criterion of substantiality, when interpreted from a realist point of view, is simply equivalent to the doctrine that an individual substance is a natural or real individual, in some sense 'given' and not 'constructed'. The term 'horse' simply *classifies* natural individuals, the horses themselves. The adjective 'brown', on the other hand, *characterizes*, not browns themselves, but brown objects. Particular browns are sliced out of reality by the concept *brown* or, to put it another way, by the nominalized predicate, 'brown' or 'brownness'. Such entities are *abstract* particulars, mere aspects of things. Their individuality is posterior to their individuation by us, or by our concepts. But the individuality of horses and the like is prior to their being classified. This is the thought (which runs so deeply through the *Essay* and the preconceptualist tradition generally) that non-substances or modes, in being logically and ontologically dependent on substances, are at the same time in some sense mind-dependent or mind-created. While individual substances are individuated by a real distinction, modes are individuated by a distinction of reason. It is that kind of mind-dependence which modern conceptualism wishes to make universal, that kind of distinction between distinctions which, in effect, it sets out to face down. In this it has been remarkably successful, if success is to be reckoned solely by the number or eminence of its adherents.

Unity

If logical priority is the most famous feature of the category of substance, the principle that substances are real or natural unities identifies what is perhaps its most fundamental feature, as well as one of the most controversial and many-sided. The notion of substantial unity has itself seemed alarmingly liable to fragmentation, and shifts in its interpretation are central to the story of seventeenth-century metaphysics.

It should by now be apparent that at least two broad approaches to the topic of unity are possible: first, through a view of a substance as the one simple subject of a multiplicity of predicates or accidents; and second, through a view of individual substances as unitary and discrete wholes, each naturally and in itself distinct in a way that abstract, non-substantial individuals are not. Admittedly the question arises whether these approaches can be presented as approaches to the same topic only by an equivocation of the term 'unity'. How, after all, can a philosophical problem about (to put it crudely) what ties together a thing and its various attributes be one side of a coin the other of which is a matter of different sorts of principles of distinction between individuals in different categories? Isn't the first problem, for example, a problem which, if it arises at all, arises equally and in the same way with respect to all subjects of predication of whatever category? What connection does it have with the question of the natural unity of horses by comparison with processions? Yet despite the *prima facie* force of these questions, it may well be that we can expect to understand predication only through understanding the most primitive or fundamental cases of it: namely, those cases in which sensible qualities are predicated of natural or 'given' subjects. Such subjects, it is arguable, are there to be picked out by sensory means prior to predication, and are to be contrasted in this respect with subjects which are picked out only through nominalized

first-order or higher-order predicates. So the topics of the subject-predicate relation and individuation may after all be rather closely connected.

In Aristotelian theory the chief work of individuation (despite the role of matter in distinguishing substances at a time) is done by the same thing as does the work of explaining the union of attributes: i.e. the form or essence which constitutes the individual and from which the properties necessarily flow. This last model for the subject-attribute relationship directly connects logic with philosophy of science, the question of logical unity with the characterization of substances as active. Such a connection remained explicit throughout the seventeenth century. As we have seen, the question of what it is for gold to 'support' great weight, a yellow colour, malleability and so forth played an important role in the arguments of Descartes, Hobbes and others just because it seemed to be a question which the 'new philosophy' answered so much better than Aristotelianism. According to the new answer, the multiplicity of sensible accidents and powers is what Hobbes called a 'diversity of seeming': in itself the one substance is simply determinable matter determinately modified. The substance-accident relation, left obscure by Aristotelian doctrine, is explained by being reduced to the perspicuous relation between 'attribute' and 'mode' (e.g. between extension and roundness) in the technical, Cartesian sense of those terms.

Here again it might appear that we are dealing with a traditional connection which can no longer have any significance for us. The present heirs or descendants of the mechanist or geometrical theory of matter would seem to be, on the one hand, current fundamental physical theory itself and, on the other hand (in so far as the two are separable), the endeavour of scientists and philosophers of science to provide an intelligible interpretation or model which makes physical sense of the mathematical content of current theory. Neither is obviously concerned with the subject-predicate or substance-accident relationship, a topic nowadays commonly felt to be more appropriately located in the narrow world of philosophers of language. Yet, although it is true that the mechanists recommended their science of body as making that logical relationship perspicuous, and so as satisfying the requirements of an adequate fundamental theory, the thought behind their claim was perhaps complex enough to suggest something still to the point. First, it involved a view of the relation between the level of experience and the level of theory. At the former level we identify sensible things and ascribe to them an indefinite multiplicity of sensible qualities and powers. The only way of making sense of this possibility is in causal terms. We have to suppose that there is something there which acts on us and on other things in a variety of ways in accordance with the laws of physics, i.e. in accordance with the fundamental nature of that 'something'. What appears as a mere multiplicity of disparate attributes in the thing calls for a common unifying explanation at the level of theory. An explanation must exist, even if we do not have access to it. Moreover, it cannot simply consist in a similar unreduced multiplicity of *theoretical* properties related only in being attributable to one thing. An adequate theory must present what is there as somehow more unitary than that, for example, the postulation of a number of distinct, disparate forces calls for a unifying explanation. More will be said in criticism of mechanistic corpuscularianism below. Yet, whatever its shortcomings, it did possess many of the features of a good theory.

To a great extent corpuscularian matter supplied the paradigm for seventeenth-century conceptions of unity, so that spirit, thought and the modes of thought were taken, implausibly enough, to bear the same sort of intelligible relations to one another as body, extension and the modes of extension. From the somewhat rickety platform provided by this analogy was announced the substantial status of spirit. Yet the modes of extension (including the modes of motion) are quantities of quantifiable attributes, whereas the modes of thought differ from one another qualitatively, in type and in content. There is no geometry or mechanics of the soul. Malebranche saw that problem, and concluded that we lack a clear and distinct idea of spirit. Leibniz, on the other hand, turned the tables on matter by finding in spirit itself, rather than in matter, the paradigmatically unitary and simple substance. Another philosopher who did the same, although by means of a very different argument, was Berkeley.

Leibniz's metaphysics involves an amazing series of linked variations on the theme of unity. There are no 'accidents' in the sense of the doctrine of predicables: *all* the attributes of a simple substance or monad are logically contained in their subject, being deducible from the complete concept of the subject. Correspondent to this logical relationship there is, as we might expect, a causal one: the essence of a Leibnizian substance is a force or law according to which the life-history of the substance unfolds. At the same time the principle that all its predicates are contained in the subject supplies the principle of individuation of the substance; any substance, actual or possible, with different attributes is a different substance. These doctrines are all versions of familiar thoughts about unity, different aspects of which are fitted together by Leibniz with the screws turned extremely tight. Combined with this abstract model, and explaining it, is Leibniz's immaterialism. The claims of bodies to be unitary or simple substances are dismissed on the grounds that every body has parts. To understand how diversity is possible in simplicity we need only to contemplate the unity of consciousness: in self-consciousness we are aware of a variety of representations in a simple subject. For a substance to have determinate attributes is for it to 'mirror' or 'perceive' the world in determinate ways. Even in coarse summary this is not the whole of Leibniz's thought about unity, e.g. 'dominant' monads can unify aggregates of subordinate monads in a relationship which he assimilated to the traditionally postulated relationship between form and matter, and saw epitomized in the relationship between soul and body.

To discuss Leibniz's exuberantly ingenious metaphysics at this point will perhaps only confirm prejudice against talk about substance, not to speak of unity. And it can be conceded that certain difficulties in his system (in particular over identity in a world without space or matter) are indications of the untenability of the notion of immaterial substance.¹⁸⁰ Yet even his thought, odd as it may seem, that the unity of consciousness can supply an understanding of the relationship between a thing and its various attributes constitutes, in effect, an intelligible response to a real problem. If the experienced variety which leads to the attribution of seemingly unconnected accidents to a body is to be explained as a 'diversity of seeming' consequent on the variety of ways the body affects observers directly or indirectly, how is the latter variety to be explained? If there is virtue in the reduction of the variety of the accidents of material things, in the name of the unity of substance, to mechanist simplicity (or its modern equivalent), shouldn't there be virtue

in a similar reduction of the variety of mental content? Leibniz simply rejected the mechanist paradigm, yet, if we find immaterialism untenable and also decline to embrace an implausibly reductive or eliminative physicalism with respect to consciousness, then a problem does arise. The present relevance of this problem lies in the possibility of expressing it as a problem about predication: what is the relation between a conscious mental state and its subject which constitutes its *belonging* to that subject? If, unlike Leibniz, we take the subject to be material, the animal, it is arguable that the only tie capable of perspicuously linking subject with predicate in this case is the identity of the state of consciousness with some physiological state of the animal. The ontological problem of consciousness is then the problem of how, in Thomas Nagel's phrase, there's something it's (subjectively) like to be in such a physiological state.¹⁸¹ But the present point is that there is nothing intolerably 'metaphysical' in approaching the issue of the relation between consciousness and the body armed with the traditional demand that the relationship between substance and attribute should be perspicuous. On the contrary, it is difficult to make ontological sense of the curious proposal, sometimes advanced, that conscious states are tacked on to their material subject simply by being 'correlated' with its physiological states. Crudely, it is not those who think in terms of the traditional notion of substance here whose views are 'queer', but those who try to do without that notion.

The notion of substances as physically unitary relates closely to the principle that substances are the ultimate, logically independent subjects of predication, and helps to explain it. If individual substances are 'given' individuals as individual modes are not, that is because (as Locke argued at some length) substances enjoy a kind of natural unity that modes cannot in principle enjoy. As it was suggested immediately above, in chapter 8, only in the category of substance are there natural or 'given' wholes or parts. We can regard a thunderstorm or a procession as a whole with parts, but every such event or process is a part of some wider event or process no less unitary or distinct. A horse, on the other hand, is a natural whole, and there is no other individual substance of which it is a part. A team of horses is not an individual substance, but a collection of individual substances. Collections, as we shall see, have members rather than true parts.

These claims are opposed to those of Frege, whose pronouncements on individuation have prophetic status for some analytic philosophers. In expounding his blanket conceptualism with respect to the natural world, Frege appealed to the fact that 'I can say with equal truth both... "Here are four companies" and "Here are 5,000 men"', but he did so without reflecting on the different ontological status of companies and men. Accordingly he regarded natural unity as a psychological matter:

The more the internal contrasts within a thing fade into insignificance by comparison with the contrasts between it and its environment, the more natural it becomes for us to regard it as a distinct object. For a thing to be 'united' means that it has a property which causes us, when we think of it, to sever it from its environment and consider it on its own.¹⁸²

For Frege, all unity is concept-relative: nothing in the physical world is individual or a

whole but thinking makes it so.

The close relationship between the question of the natural or physical unity of individual substances and the question of the logical unity of substance and accidents is strikingly, even spectacularly illustrated by the topic of compound or composite terms, introduced in chapter 2 and discussed at various other places above. The possibility that such terms reflect the peculiar logical status of the category of substance was hinted at in the table of natural and non-natural substances and modes in chapter 8: no modes appear under the column headed 'compounds'. Why?

One line of approach to this question is through the Aristotelian view that the only true species are species of substances. The general doctrine was alluded to only briefly by Locke, perhaps because he could not endorse it:

in *mixed Modes* any of the *Ideas* that make the Composition of the complex one, being left out, or changed, it is allowed to be another thing, *i.e.* to be of another species, as is plain in *Chance-medly, Man-slaughter, Murther, Parricide*, etc.

183

To put the point in reverse, if we define a new term by adding arbitrarily to the notion of a species of mode, we create a new species; but not if our definition adds to the notion of a species of substance. If we suppose *thunderstorm* to be introduced as a new concept, equivalent to *storm with thunder*, then thunderstorms would have no less right to be regarded as constituting a species than the class of storms. But the concept *thunderstorm* could as easily be divided in turn. Thus there are no 'last species' of modes. Yet, if 'baker' is defined as *man who bakes bread*, or 'adult' as *fully-grown animal*, the class of bakers or of adults does not constitute a species. The parts of the concept *baker* or *adult* do not, as it were, amalgamate: *baker* remains irremediably a composite concept incorporating, but not uniting, species plus accident.

Approached from this direction, the contrast needs to be drawn with some care and qualification, not least because of the doubtful objectivity or reality of the species and, still more, the genera of substances. Locke, of course, was attacking the doctrine from the point of view of his scepticism about any natural species at all. But the Aristotelian tradition fortunately supplies the material for another, rather less cluttered approach, through the topic of individuation.

An adult is a fully-grown animal, but the human adult is identical with the boy or girl who grew up; that is, the adult began to exist when the man began to exist, not when the man became fully-grown. A baker is, by definition, a man (*homo*, human being) who bakes bread for a living, but that addition to *man* has no effect with respect to the identity-conditions or existence-conditions of a baker. A baker who retires or changes occupation does not thereby cease to exist. By contrast, although 'downpour' may be defined as *heavy rainfall*, in particular cases the downpour and the rainfall may well be distinct, since the latter may have been taking place for hours before the downfall began (*i.e.* before the rainfall became heavy), and may continue to do so after the downpour is over. Similarly the place of the rainfall may be larger than the place of the downpour. As we might put it, although 'downpour' is defined as *heavy rainfall*, the downpour is not

the rainfall but the heavy-rainfall: the parts of the definition amalgamate. Hence it is only if the rainfall is always and everywhere heavy that it is identical with the downpour. Otherwise the downpour is a part, temporal or spatial or spatio-temporal, of the rainfall.

Recent writers on identity are not unaware of the class of composite terms. One of the more sensitive to these issues, David Wiggins, seems first to have noticed them in the case of what he called 'phase-sortals'. Since then he has made further use of the notion, introduced there, of concepts which are 'restrictions' of more general underlying sortal concepts.¹⁸⁴ In these, he says, 'the principle of individuation is certainly present, but it is buried and must be disinterred' from what is irrelevant to it. The principle of individuation associated with (his examples) 'boy' and 'Anzac' is supplied by the 'buried' concept *man*. Whether or not the notion of a 'restricted' sortal offers an adequate account of composite terms (and we shall see that it does not) it should be recognized that the bare existence of Wiggins' examples presents a twofold problem of fundamental importance to the theory of identity. First, *why do clear instances of such terms exist only among noun-predicates predicable of substances* (or perhaps, more generally, of substance-like continuants)? And second, *what constraint here prevents us from manufacturing individuals by arbitrarily tinkering with principles of individuation?* In other words, what prevents us from introducing a substance-term, say 'bakist', to express the concept *man who bakes*, while yet intending this definition in such a way that the sense of the term *does* contain a principle of individuation distinct from the principle associated with the term 'man'? What prevents the parts of the concept from amalgamating or uniting in the same way as the parts of the concept *downpour*?

Now these questions must either be rejected on adequate grounds, or it must be accepted that it is a shortcoming of current theory of identity that absolutely nothing is done to answer them. To announce that the concept *man* is individuable *per se*, and that the concept *baker* is individuable only in virtue of its including or restricting or entombing the concept *man*, whereas the concepts *rainfall* and *downpour* are both individuable *per se*, is of course to give no answer. It is merely to give a redescription of the problem in terms favourable to conceptualism, i.e. consonant with the assumption that 'concepts' of some special kind are responsible for slicing out individuals.

The best bet for the conceptualist might therefore seem to be to reject the presuppositions of our two questions. But can they be rejected? Well, with respect to the first question it might be argued that it is only through a kind of linguistic accident that there seem to be no clear analogues for 'girl', 'bachelor' and 'baker' among nouns predicable of modes. Suppose it to be stipulated that 'riot' means *disorderly assembly*. Then the time at which the riot began (it might be conceded) would be correctly taken by English-speakers to be the time at which the assembly became disorderly. Yet why (it might be asked) should we not understand the definition in such a way that the riot *is* the assembly, and began when the assembly began even if the assembly began and long continued as a peaceful demonstration? The rough answer seems to be that it is impossible to make sense of the suggestion that a riot was going on when nothing satisfied the predicate 'riot', i.e. nobody was rioting. By contrast, the parallel claims that the baker existed before anything satisfied the predicate 'baker', and that the caterpillar continues to exist after it has ceased to satisfy 'caterpillar', are logically impeccable. Note

that nothing is gained by stipulatively defining a term (e.g.) ‘orgy’ in some such way as *party at some time in the course of which such and such a degree of impropriety occurs*. That definition might achieve by force of stipulation an unnatural answer to the question ‘When did the orgy begin?’, but it would not achieve an analogue to ‘baker’. For under it a party which begins with the guests’ respectably sipping sherry but ends with large-scale impropriety over the Sauternes would satisfy the predicate ‘orgy’ *throughout* its existence, whereas someone satisfies the predicate ‘baker’ only while their occupation is the appropriate one.

In order to reject the presupposition of our second question, on the other hand, it would be necessary to deny, in the face of the logical phenomena, that there are any constraints on the manufacture of principles or criteria of identity for substances. That is the response to be expected from the mainstream conceptualist, for it is an unargued and unquestioned assumption by many writers on identity that we can arbitrarily stipulate modifications of ‘criteria of identity’ associated with substance-terms, just as (they suppose) we can stipulate ‘criteria of application’. But if there is such a thing as an obviously false claim in philosophy, then that is such a claim.

It remains to consider such connections with other attributes of the category of substance as may help us to understand the phenomenon of composite terms. It might be argued that their possibility is simply an aspect of the duration of substances, i.e. of substances’ lacking temporal parts and existing all at once. For it is only in the context of the false doctrine that a substantial object has, or may be regarded as having, temporal parts that it can be at all plausible that such predicates as ‘baker’ or ‘boy’ pick out individuals distinct from the man. For these supposed individuals, it can then appear, share matter with the man as unproblematically as his hands and fingers share matter with the man. Moreover, there are non-substantial ‘continuants’, such as nations and whirlpools, which have borrowed the logic of substances to the extent that they exist all at once and may thereby seem as capable as substances of sustaining composite predicates. The possibility that a nation should be for a time a debtor or a creditor does not prompt us to distinguish the debtor from the nation.

Nevertheless there are reasons for rejecting the conclusion that a full explanation of composite terms is to be found in the duration of substances or, more generally, of ‘continuants’. First, there is the point that the spatial, as well as temporal extent of individuals is in question. Second, in the case even of the most substance-like non-substantial continuants, the logical constituents of their quasi-composite predicates tend to amalgamate as the constituents of composite predicates never do. A kingdom is a nation or state ruled by a king, but we may nevertheless suppose that the kingdom of Italy ceased to exist, as the Italian state did not, with the formation of the republic. It is true that we may prefer to identify the kingdom with the republic, and both with the nation or state. But there is no requirement to do so, as there is to identify the baker with the man or, if he changes his job, with the butcher. The continuity of the nation is no more ‘given’, or less notional, than the extinction of the kingdom. To a nationalist, the nation may seem the more natural individual, to a monarchist, the kingdom. But that issue is political, not ontological. So too with non-substantial continuants less dependent for their existence on human convention. Suppose that ‘hailstorm’ is stipulatively defined as *storm*

with hail. Suppose too that both storms and hailstones are conceived of, not as events, but (with hurricanes and tornados) as meteorological disturbances which move from place to place and undergo change, i.e. as substance-like continuants. Nevertheless there would be no constraint on allowing that a particular hailstorm has had a shorter life and narrower extent than the more general storm of which for a time it has formed a part.

From such examples as these it seems clear that the amalgamation of the parts of a definition reflects the notional status of the individuals denoted by the terms defined, while the logical phenomenon of compounds or Aristotelian *syntheta* derives from the status of substances as natural or 'given' individuals. For non-substances or modes, as we have seen, classification is prior to individuation and (roughly) is relative to a predicate predicable of something else. Particular modes are abstractions sliced out of reality by general concepts. It is therefore unsurprising that modification of the predicate is liable to breed fresh individuals. Particular substances, however, are naturally discrete individuals whose given individuality is prior to their classification by us. It is therefore unsurprising that mere addition to the classificatory predicate we employ will *fail* to breed fresh individuals.

How, then, is the natural or 'given' individuality and unity of substances to be explained? In chapter 8 the proposal that substances are united by a principle of causality or activity was seen as insufficient by itself to distinguish the real or natural unity of individual substances from the sort of unity possessed by some natural processes. Before a better explanation is proposed, it would be appropriate to consider the other features of the category of substance on our list.

Materiality

The doctrine that substances are compounds of form and matter was not based on nothing. Paradigm substances drawn from the ordinary objects of experience are indeed organized matter. Exemplary immaterial substances (the Prime Mover, the active intellect, the Christian God, immortal souls, the pure ego) are all metaphysically queer, beset with difficulties in their conception. It will therefore be assumed that notions of immaterial substance are aberrant and incoherent, and that all substances are material objects.¹⁸⁵

The materiality of substances is evinced in the division of the category of substance into 'things' such as dogs or oak-trees, and homoeomerous substances or stuffs, such as gold or water. This is a division of types of substance corresponding to which there are two types of individual: on the one hand, 'things' and, on the other, such quantities of stuff as the gold in this ring or all the tea in China. The latter are perhaps dubious 'individuals'; but, however that may be, they are certainly not 'given' as such, since the boundary of whatever quantity of stuff we wish to identify is up to us.

Modern conceptualist doctrine is ambivalent in its attitude towards materiality. On the one hand most conceptualists are prepared to talk without qualms of a category of 'material objects' and even to allow it logical primacy within the conceptual scheme we have (although for some the honour is shared with events, or even with numbers). On the

other hand, formal explanations of the thing-stuff dichotomy are commonly advanced at a level of abstraction which carries it up out of the category of substance altogether. So-called 'count-nouns', among which are 'walk' and 'laugh' as well as 'horse', are opposed to 'mass-terms', which include 'walking' and 'laughter' as well as 'snow'. If so highly abstract an approach were successful, i.e. if it could give us an explanatory account of the thing-stuff distinction, then such entities as laughs and battles would have every right to be regarded as material objects. Battles would then seem to be composed of fighting, or of such more specific stuffs as advancing and retreating. Walking, it could be held, is the matter of walks. Yet that is so, as it seems obvious, only metaphorically or analogically. The logical analogy is, of course, imperfect: the gold which once composed a ring can be used again, but not the laughter which composed a laugh.

It is substances alone which are literally or primarily material. Something like the thing-stuff distinction, and so something like materiality, can be extended analogically to other categories in the process of nominalization. In nominalization we, as it were, create or construct entities. As the eighteenth-century writer on language, James Harris, put it, we 'convert even attributes into substances',¹⁸⁶ and in so doing we have a choice between things and stuffs. It is often convenient to derive both sorts of noun from the same predicate, as it is convenient both to speak of the unbounded activity of walking, and to impose bounds on stretches of that activity, *walks*. So the distinction between 'count-nouns' and 'mass-terms', as it applies to such abstract beings as walking and walks, needs to be explained by reference to the less general paradigm, the distinction between substantial things and stuffs. The order of explanation here is one aspect of the primacy of substance. Unless there were natural individuals and stuffs, there could be no constructed ones.

The traditional view that individual substances are distinguished at any one time by their matter can be expressed in two equivalent ways: first, as the principle that two distinct individual substances cannot be composed of the same matter at the same time and, second, as the principle that if two individual substances have at any time all parts in common they are identical. Modern philosophers often deny the former and radically reinterpret the latter by introducing the paradoxical notion that material objects have temporal parts. So much the worse, perhaps, for modern philosophy.

Duration

On the traditional view, as for common sense and some modern philosophers, a substantial thing exists all at once, occupying space but not time. It persists through time and change, gaining and losing material parts. An event or process, on the other hand, unless strictly momentary (if such a thing is possible), takes time to exist. It does not undergo change but, as it were, unfolds. It has successive temporal parts which cannot, of course, be gained or lost.

Nevertheless it may be that today more philosophers than not hold that there is nothing seriously mistaken, even if something is linguistically unusual, in treating things as four-dimensional time-sausages with temporal parts, 'stages' or 'phases'. A primary motive

for the modern paradox has no doubt been the attraction of a logic supposedly more consonant with the theory of relativity, but for many the notion of a stage is justified by its value to the theory of identity. There are some dissentient voices, however, and David Wiggins rightly finds a 'terrible absurdity' in the claim that a material object is a long event.¹⁸⁷ Yet he does not pause to justify his judgement, for example by exploring the ways in which the absurdity infects views about identity. Another way of justifying it is to draw attention to the mutually explanatory relationship between the *endurance* of substances and other properties of the category, other grounds of the distinction between substances and events. For example, it goes without saying that the peculiar capacity to lose old parts and gain new ones is a function of the literal *materiality* of substances.

Activity

Substances are accorded a special role by the tradition in causal explanation.¹⁸⁸ They are not, as we have seen, the only familiar or pre-theoretical entities with 'natures' to which we can appeal for the purposes of explanation. An occurrence might well be explained by reference to the nature of some mode or other, e.g. measles or corrosion. Nevertheless the 'nature' of gold or a horse has a different status from that of the 'nature' of a disease or chemical process. Here there is an intuitive connection both with the materiality and the endurance of substance. For in them lies the possibility of natures which endure *through* the changes and interactions which they explain, natures embodied in the structure of constituent matter.

It is to make effectively the same point to say that substances possess *powers* as modes do not. We can, it is true, make sense of certain sorts of ascription of power to non-substances. A voice, for example, might have the power to instil fear; or a sound, the capacity to break wine-glasses. Yet a voice or sound, in this context, is not something fully particular, like a piece of gold or a particular utterance. A particular speech-episode either frightens or not, and there seems little point in attributing to an episode an unexercised potentiality simply on the grounds that in different circumstances it would have had a different effect. Better candidates for the status of subjects of dispositions can be drawn from just those modes which are logically most analogous to substances: waves, storms and the like. Yet, however that may be, it seems certain, for reasons given in chapter 8, above, that in order to suppose a full-blooded power it is necessary to suppose an enduring thing with a (possibly unknown) structure, capable, in virtue of that structure, of affecting other things in an indefinite variety of ways and of suffering an indefinite variety of changes itself. And that is to suppose a substantial thing.

One objection to according substances their traditional pride of place takes the form of an argument that the explanation schema embodied in the notion of a substance is outmoded. When in the mechanist youth of modern science the notion was reinterpreted, the essences of such things as horses and gold were no longer thought to be ultimates; but the new ultimate, matter with an essence, seemed to fall squarely enough into the category of substance. No doubt it was tailored to do so. The ultimates of present-day physical theory, however, no longer seem to do so. Electrons, photons and the like are not

helpfully conceived of as shaped space-occupants composed of a uniform superstuff. It may be that quantum mechanics, as much as the theory of relativity, has fostered the tendency to underplay the logical differences between things and events. Physics can seem to have shown that material things are really probabilistic processes.

If the entities postulated by physics cannot be straightforwardly categorized as 'substances' or 'material objects' (i.e. if the standard properties of the category of substance do not all pertain to them) that is not, on reflection, surprising. It should not at any rate surprise those prepared to make a distinction which is advocated in the present work on other grounds, but which conceptualists are almost bound to reject. That is the distinction between the pre-theoretical level of experience and the level of speculative theory about the fundamental nature of the objects of experience. Given that distinction, then it is very intelligible that the interlocking features of the category of substance, especially in so far as they involve the natural and real individuality of substances, reflect and signal the impingement of reality upon us (unmediated by any 'concept') at the level of *experience*. At what other level *could* there be direct cognitive contact with reality? *Nature, thing, structure, power, stuff*—all these notions function in their relation to the category of substance in such a way as to mark what is properly assumed to be open to a full theoretical account or explanation, and to mark it in a way which is neutral as between rival accounts precisely because this, their primary employment, is at the pre-theoretical level of experience. If, as seems inevitable, the physical theory employed in that further explanation involves the postulation of further individuals beyond the limits of direct experience (such as the primary particles of modern physics), then it is clearly impossible, *ex hypothesi*, that those individuals should be 'given' in the same way as complex and sensible material objects: i.e. that we should ever identify them without knowing what kinds of things they are, and without having some systematic, at least approximate, knowledge of their nature.¹⁸⁹ For theoretical entities are individuated *through* our speculation as to their nature and kind, as perceived material objects are not. It is therefore unsurprising that such objects do not fall into the category of substance as traditionally described. So far is this from being an argument against realism, it is just what the realist should expect. Which is not to say that there is nothing problematic about the ontology of quantum physics, or that there are *no* constraints on what makes sense in a description of reality at that level.

This issue surely identifies an important source of the crisis of confidence in the category of substance which led to its present disrepute. In Aristotelian philosophy the theory of substance was developed in the context of the assumption that, for the most part, scientific explanation and theory could effectively begin and end with the natures of the relatively large and complex individuals, above all biological individuals, which we can pick out prior to theory at the level of experience. When it came to be felt that a science based on that assumption left too much unexplained, the mechanical hypothesis offered an alternative model which still for a time held the levels of experience and theory together, although in a different way. For essentially the same mechanical concepts and the same category of a *stuff* as we employ at the level of experience were employed in a theory about the imperceptibly small. Indeed, it was evidently felt necessary to show that matter subject to mechanical modifications fitted the traditional

philosophical conception of substance even better than the sensible material objects in relation to which that conception had originally been constructed. Other felt needs, relating to God and the soul, played a distorting role in the debate. But the real crisis came with the fall of mechanism and the sense that the traditional conception of substance is inappropriate, at least as it stands, for dealing with the postulated entities of physics. In effect the theory of substance was split between the philosophy of physics and the philosophy of experience (or of the way we think about the world at the level of experience). It is not, perhaps, entirely surprising if some have felt that the second fragment is a disreputable and reactionary relative of the first fragment, a relic of outdated world-views. Yet that judgement is insensitive to the possibility of, and need for, a philosophical explanation of the primitive structure of our thinking: in particular, an explanation of the contribution made to that structure by reality itself, not through its tendency to confirm or disconfirm the explanatory stories scientists tell about it, but by its impingement on us at the pre-theoretical level at which we make our first and most fundamental discriminations.

Membership of a, natural kind

The still generally accepted principle that every individual substance belongs to a universal kind or sort has had, in its history, two very different sources of support. First, together with the view that the only natural and real species and genera are the species and genera of substances, the principle constituted part of the metaphysical conception of individual substances as embodied forms. The second grounding of the principle, on the other hand, has been the conceptualist theory of identity which began to be formulated by seventeenth-century opponents of Aristotelianism, most notably by Locke himself.¹⁹⁰ The universal kind is held to be essential to the individual substance just because a universal idea or concept must come before individuation: individuation is always relative to a sortal concept.

In its standard modern form, however, the conceptualist theory takes *automobile* or even *piece of stone* to be as good a sortal concept as *horse*. If it is thought that all individuation is a matter of our slicing up the world by means of concepts, there might well seem no basis for the careful Aristotelian distinctions between the natural and real unity of a horse, the accidental unity of a piece of stone or a dead body, and the artificial unity of an axe. On the other hand, if we are neither Aristotelians nor conceptualists, it may well seem that a lump of lead is quite an impressive individual in its own right, perhaps no less so than a plant; and that artifacts, while sometimes physically unitary, like most teaspoons, are sometimes hardly more than notionally so, like a steam-engine. In that case, since 'lumps' or 'pieces' do not form a kind in any significant sense, the principle that individual substances belong to kinds would seem dispensable. The converse principle, however, that the only real or natural species are species of substances, still has much to be said for it, some of which has been said above under the heading 'Unity'. These issues will be a topic of Part III.

Conclusion

It is to be hoped that the above discussion of six properties traditionally ascribed to the category of substance has presented the theory of substance (or something distilled from it) as more plausible, more unitary and coherent, and more capable of explaining the actual structure of our thought and its relation to reality than common prejudice allows. Yet it leaves our enquiry more or less at the place reached at the end of the previous chapter, if perhaps better equipped to proceed. The crucial question of the natural individuality or unity of paradigm substances will be a chief topic of Part III, but there have been heavy enough hints of the general form which the answer will take. The explanation of that unity lies with a further connection between the various conditions of substantiality.

Substances, and only substances, are material. Only substances, therefore, can have the kind of given physical boundary that a coherent material thing can have. Only substances can be *materially* unitary and discrete objects. This somewhat obvious answer has not been popular, but it does have its antecedents in the writings of those ancient and early-modern atomists (including certain arguments in the first edition of the *Essay*) for whom in effect the continued existence of an individual substance just is the continued existence of a coherent mass of particles or matter. For such an answer to be adequate, however, it is necessary to say more about what constitutes a natural individual's 'coherence'. It is implausible to suggest that a cannon-ball and the piece of chewing-gum sticking to it together form such an individual. What matters is not just coherence, but also the nature and causality of the coherence. Thus the notion that substantial unity is constituted by a causal principle or principle of activity is not entirely mistaken, but it needs to be subordinated to a conception of material unity. It will be time to take the argument further when Locke's own developed theory of identity is under consideration below.

Part II
God, Nature and the Law of
Nature

10 Introduction to Part II

In the opening paragraph of the early manuscript now known as the *Essays on the Law of Nature*, Locke endorsed an ancient analogy. A divinity presides over the world by whose command there have been set for all natural objects, from the heavenly bodies to living things, laws appropriate to their natures. Can we then suppose that man alone is exempt from law? There is such a law for man, Locke asserted: the moral law or Law of Nature.¹

One antecedent of this analogy is Aristotle's conception of function as the foundation both of natural science and of ethics. The essential function of man both explains what man does, as if biologically, and what man ought to do, the good for man. But for Christian theology the analogy possessed other implications. These retained a certain force throughout Locke's later philosophizing, despite both his rejection of Aristotelianism and his own efforts to stress and to explain the differences between ethics and natural science in respect of their objects and epistemological status. Hence it is appropriate and even, perhaps, obligatory for a commentator on the *Essay* to discuss physical and moral laws at least sometimes in the same breath. Their conjunction in the chapters following is not intended to encourage goggling at Locke's antiquated otherness, but, more practically, to facilitate the interpretation of what he wrote on each of these distinct topics. The crucial common question concerns the depth and character of Locke's 'voluntarism': crudely, are his laws of nature, in whichever sense, arbitrary and contingent dictates of God's will, or are they necessary, flowing from the natures or essences of their objects? Could, for example, just this substance, matter or body, have been created subject to different mechanical laws from those which actually obtain, or did the act of creation, in determining which substance should exist, necessarily and *ipso facto* determine which laws it should obey? Equally, could God have willed a different moral law for rational, sensitive creatures from the one he has in fact willed?

Although their treatment in any detail lies beyond the scope of this book (and its author's competence), it may nevertheless be helpful here to gesture, however inexpertly, towards some of the theological issues which were entangled with the above questions in seventeenth-century thought.² First, there was wide and long-standing, although not universal, agreement that nothing uncreated could exist besides God. Consequently neither moral universals nor the essences of natural kinds were allowed to have existence independent of God. Yet there were very different conceptions of the nature of their dependence. One question was whether universal archetypes in God's mind were the kind of thing that needed to be created. Aquinas had held that God is himself bound by Natural Law, even though it is dependent on his reason for its existence. In this tradition, Suarez argued that eternal truths are merely hypothetical, so that essences, although not nothing, are conceptual entities independent of God's will for such existence as they possess. They

do not need to be created. For Ockham, on the other hand, God's omnipotence and freedom are absolute: he is not bound by any antecedent or eternal universal thought to any particular act. As the creator of good and evil, he perpetually lies beyond them. Natural Law is simply the contingent order devised by God. Consonant with this view was Ockham's strong nominalism and rejection of universal natures.

With this difference went another. Roughly, if there is something intrinsically rational and universal in Natural Law, then there is the possibility of our coming to know it through the employment of natural reason. Indeed, it could be said that we possess natural reason just in so far as there is a certain correspondence between our intellect and the divine intellect. On the other hand, if God's laws are absolutely arbitrary (in effect, he is free to change his mind), then we should not expect to discover their content by the employment of natural reason: in the case of physical law, experience will inform us, so far as it can, of brute regularities; in the case of moral law, we may recognize which acts are beneficial to mankind, and innate conscience may be presumed to point towards what is right, but a more promising and authoritative route than either of these to God's arbitrary dictates would seem to be revelation.³

The notion of absolutely untrammelled omnipotence figured in another controversy standing behind much seventeenth-century theology. The distinction between the elect and the damned was regarded as a matter of God's arbitrary decision. Augustine's doctrine of the original corruption of Adam's descendants after the Fall made 'prevenient' or comprehensive grace a necessary condition of the power to choose and pursue the good. Without grace we have freedom only to sin. For Pelagius, as for the pre-Augustinian tradition, the Fall had a different implication: man's moral nature has been weakened, but we are free to choose between good and evil, and may by our choice merit auxiliary grace, God's help in overcoming the evil in our nature. That is to say, grace can, for Pelagianism, be earned by the effort to be good. On an Augustinian view, Natural Law may exist and be knowable as a measure of sin, but is more or less irrelevant to salvation, since we are motivated either by corrupt desires or by grace. On a Pelagian view, our recognition of Natural Law can supply a rule of conduct, a means of grace open and visible to all. The former had been adopted in essentials by Luther and reformed theology. Something like the latter had been brought against Puritan critics of the Church of England in the late sixteenth century by Hooker, who argued that church order, like the laws of a state, could be founded directly on Natural Law as well as on revelation. Hooker was more than once an object of Locke's express approval.

All philosophers of Locke's time, even if not (as many were) strongly and independently motivated by theological concerns, would have been acutely aware of the theological minefield through which they would need to pick a way in giving an account of physical laws, moral principles or our knowledge of either. The consequence is that in this area careful equivocation, measured concessions to philosophically awkward but revered doctrines, and unobvious reinterpretations of what had to be retained in the letter, if not in the spirit, may cause more than normal problems for the reading of a philosophical text. Quite apart from the crude motive of self-preservation (and Locke was evidently concerned to avoid irksome and damaging controversy with theologians, for example about the Trinity), such expedients were rhetorically necessary for a Christian

writing to convince Christians. Because of the standing of scripture, the creeds, the thirty-nine articles and other authoritative writings, it was endemic to the discussion of certain issues that dispute should have been less over sentences than over their interpretation, while principles most naturally understood as opposed might seem to be embraced by the same individual. Notice was taken in Volume I of just such theologically motivated equivocation in the *Essay* in Locke's apparent acceptance of immediate revelation as a source of infallible knowledge just at the point when the main thrust and purpose of his argument, without any doubt or much disguise, was to reduce its deliverances to the level of probable belief at best. Another example discussed in Volume I was the contrast between his relatively innocuous published remarks about miracles and the carefully worked out thesis of the unorthodox *Discourse of Miracles*, unpublished in his lifetime, which is so much more obviously consonant with his main philosophical principles. Partly for reasons of the same sort, it is argued below, his assertions relating to the status of natural laws can seem equivocal or even contradictory unless due care is taken.

With respect to one theological question, at least, the implications of Locke's position in the *Essay* and elsewhere seem clear enough. He held that the moral law can be known demonstratively, by the natural light, and that we are free to act on our knowledge without extraordinary grace. As in the case of the Remonstrants who befriended him in Amsterdam, a preference for a minimum of dogma went with a preference for the notion of salvation by works over that of salvation by faith. The interpretive difficulty arises rather with respect to his combination of such rationalism with the assignment of Natural Law to God's will. If Natural Law can be known, and so is necessary, how can it also be arbitrary? The question arises equally and, as far as concerns the interpretation of the text, even more insistently with regard to physical law. Although Locke placed knowledge of 'the original Rules and Communication of Motion' beyond human reach, for at least some of the time (e.g. in the arguments relating to our ignorance of substance and essence) he represented them as, like the axioms of geometry, necessary and intelligible in themselves. Our ignorance is due to the weakness of our faculties, rather than to the contingency of the rules. Yet at the same time 'we cannot but ascribe' those rules 'to the arbitrary Will and good Pleasure of the Wise Architect'.⁴ How then are these different voices to be reconciled?

That is the chief question to be attacked in the next two chapters, a discussion to be followed by an assessment of philosophical mechanism. The argument will then move on to a consideration of Locke's proof of God's existence, and so to his demonstrative theological ethics and the analysis of human freedom of which he was particularly proud. In the concluding chapter some morals about ethics will be drawn.

Forms of mechanism before Locke

'Mechanism' is a name that can be given to a variety of philosophical positions, but it can provisionally and roughly be defined as the view that the perceptible functioning of machinery supplies an overt illustration of the intelligible principles which covertly govern nature as a whole. According to the purest form of mechanism,⁵ the understanding which is in principle possible of mechanical processes is the same in kind as the understanding which can be achieved in geometry. Knowledge of the actual structure of the machine should enable us to understand why it must operate as it does. For what I will call 'pure' mechanism the 'must' is a hard or geometrical 'must', since otherwise there is something in the operation of the machine which is not fully perspicuous and intelligible. The intuition underlying such a view is expressed with clarity in the Port Royal *Logic*, despite its Cartesian sympathies. If someone observes that an appropriately shaped axle passes through two millstones, the upper millstone having a square hole through it (where the axle is square) and the lower a round hole (where the axle is round), then he can employ reason to predict with certainty that the axle will turn the upper stone without turning the lower one: *'l'effet qu'il a prétendu s'ensuit infailliblement'*.⁶

Descartes was in the *avant garde* with respect to the geometrization of nature, but he was not an advocate of mechanism in its purest form, in the present sense of 'pure'. It is true that Cartesian matter, considered in itself, has nothing but mechanical or geometrical attributes, extension and motion and their modes. Yet the laws of mechanics do not, as it seems, stand to reason on their own, but need to be derived from the activity and immutability of God. God, by an action equivalent to continuous recreation, maintains both matter and the overall quantity of motion and rest 'by his normal participation'. From this principle are derived the laws of motion, described as the 'secondary causes' of particular movements.⁷

The first Cartesian law is that every body, *'quantum in se est'*, remains in the same state, whether of motion or rest, until some external cause impedes or propels it. The second law is that, of itself, movement is in straight lines. The third law concerns impact, which Descartes saw in terms of a conflict of forces, a force to move and a force to resist motion, the latter being a distinct force of rest. In general the 'stronger' body transfers some of its motion to the 'weaker'. The rules for the application of this law to specific cases distinguish the opposition between movement and rest in the two bodies from the opposition between the determination of one body to move in a given direction and the power of the other body to alter that direction according to the direction of its own movement.

Further details of Descartes' theory do not concern the present argument. One question

which does, however, is this. Was there any motive stemming from Descartes' mechanics itself for his deducing the laws of nature from God's immutability rather than treating them as in themselves axiomatic? Is it perhaps that God was called in to ground a notion of force? Taking the first law in isolation, it is not easy to see any reason unless a theological one for the derivation, since other mechanists treated the law as evident in itself. Descartes' friend Isaac Beeckman had stated in 1614 that, rather than suppose that a force is implanted in a thrown stone, 'it is easier to conceive that a moved body in a vacuum will never come to rest, because no cause of change meets with it: nothing changes without there being some cause of the change'.⁸ Far from attributing a force to matter over and above its geometrical properties, the first law might well seem to be an expression of the very inertness of matter, and a principle which stands to reason in its own right.

Much the same goes for the second law, but the third law and subordinate rules of impact have seemed to some commentators to introduce calculable physical forces incapable of being reduced to the purely geometrical modes of matter which are their effects. It has been argued that Descartes' conception of the force of rest in particular is irreducibly dynamic.⁹ Here then is a possible motive for the role ascribed to God, if we suppose that Descartes felt *both* that to indulge in the notion of forces as inherent modes of matter is to revert to the postulation of the non-explanatory powers and occult qualities of Scholastic philosophy *and* that a purely geometrical idea of matter in motion is incapable of generating a systematic mechanics. Indeed, rather more importantly, he was likely to have felt that to do the former is to offer an account contrary to the proper understanding of what it is to be a 'mode'. This last point is important because, as we have seen, the more fundamental objection to unreduced powers from the point of view of metaphysics was not the easily grasped complaint that they are pseudo-explanations, but the now less well understood objection that their purported relation to the substances in which they are supposed to inhere is simply unintelligible. By contrast, the relation between any determinate shape or motion and the extended substance it modifies was taken to be wholly perspicuous. How could any extended substance exist *without* determinate limits or determinate motion or rest?

Some support for such a hypothesis as to Descartes' motives for introducing God might be drawn from comparisons with Hobbes, a 'pure' mechanist who expressly reduced power, endeavour, resistance, impetus and force to motion. As Hobbes put his view, 'power is not a certain accident which differs from all acts, but is, indeed an act, namely motion, which is therefore called a power, because another act shall be produced by it afterwards'. Hobbes was explicit (presumably against Descartes) that 'rest does nothing at all', and that 'nothing but motion gives motion to such things as be at rest, and takes it from things moved'.¹⁰ Nevertheless the attempt to find in the role ascribed to God by Descartes a response to the special requirements of his concept of force, and in particular the force of rest, faces considerable difficulties. For, whatever modern historians of science may make of these requirements, Descartes himself expressly stated that force or power consists 'simply in the fact that everything tends, so far as it can, to persist in the same state, as laid down in our first law'.¹¹ If we take this statement at its face-value, it seems that the question of Descartes' motive for deriving his physics from

his metaphysics rests wholly on the status and derivation of the first law. We are therefore left to presume that his only real reasons for his emphasis on the first cause were indeed theological.

Descartes clearly did have a theological motive, in that it was important to him that the laws of physics flow from God's nature, i.e. from his will determined by that nature, rather than simply from the nature of the substance which he has arbitrarily chosen to create. At the same time Descartes notoriously held that even the most elaborate details of creation, the system of heavenly bodies, the species of animals and plants and so forth, have flowed (or, more discreetly, *could have* flowed) mechanically from unorganized matter governed by those general laws, rather than from any particular divine design or arbitrary manipulation of matter.¹² In short, Descartes preferred the conception of God as the necessary ground of all being and all happening to the model of the divine architect and craftsman.

Two further considerations relating to Descartes' mechanics should, however, be mentioned. First, it seems that the role ascribed to God did enable Descartes to employ certain kinds of explanation in his derivation of the specific rules of impact which might otherwise not have been open to him. The third rule, for example, is justified by the consideration that it is the easiest or simplest possibility. As others have pointed out, this is in effect a teleological argument, although the principle of divine economy involved is very different from the Aristotelian teleology which Descartes so strenuously rejected.¹³ More significantly for present purposes, God's role might seem to place the whole phenomenon of the transference of motion in a light more agreeable to Descartes' metaphysical principles. Motion, as a mode, is tied to its substance, so that talk of its 'transference' from one body to another was for Descartes at best a sloppy and metaphorical description of the phenomena.¹⁴ God's agency in conserving the overall quantity of motion is the reality behind the metaphor.

Whether or not the principle of economy and the notion of transference did supply important motives for the 'impurity' of Descartes' mechanism, it may advance my present purpose to consider a further question of interpretation. This question is raised by the suggestion that determinate forces were after all seen by Descartes as genuine attributes of matter, correlates of divine agency 'immanent in "nature" or extension'. Gueroult advanced this interpretation in a well-known discussion in which he claimed that the forces of rest and motion are 'in nature' as "second causes" governed by laws that the immutability of the Divine will grounds according to the three principles of conservation, inertia and the rectilinearity of motion'. Yet Gueroult's argument appears stretched and uncertain at crucial points. For example, he relied on a doubtful interpretation of Descartes' statement in a letter that 'moving force...in created substance is its mode, but it is not a mode in God'. Gueroult's argument requires 'created substance' here to include body, whereas the context indicates that it was only finite spirit which was in question. Indeed, Descartes had just stated that, from the point of view of general mechanics, 'moving force is the force of God Himself'.¹⁵ It may therefore be best to assume that Descartes regarded talk of force in bodies, like talk of the transference of motion from one body to another, as a popular misrepresentation of the metaphysical reality.¹⁶

It will be seen that there are nevertheless three forms of mechanism each of which might be ascribed to Descartes with at least some plausibility. First, it can appear that the 'impurity' of his mechanism was essentially based on theological considerations, and had little relevance to the actual form of his mechanics. On this interpretation, he was advancing what was in effect a 'pure' mechanism, but from within a certain theological framework. Second, it can be argued that he relied on God's role in order to deal with difficulties in his physical theory, such as the force of rest or the transference of motion. Third, as an elaboration of this latter view, it has been held that force is on Descartes' account genuinely a 'mode' of bodies, if one which exists only in reciprocal relation to God's activity. A fourth view, of course, might take his theory to be indeterminate as between these possibilities. Interpreted in the second and third ways, Descartes was committed to 'impure' mechanism and a correspondingly strong voluntarism. Interpreted in the first way, he took the laws of motion to be as self-intelligible as the axioms of geometry, while God's continuous action in maintaining them in being would seem no different in kind from his action in creating and maintaining the other eternal truths.

Whatever its correct interpretation, Descartes' theory loomed large in later seventeenth-century philosophy of mechanics. Malebranche, who was closely studied and criticized by Locke, was perhaps the most famous advocate of the view that natural causality is a fallacious appearance of what is in fact the direct agency of an omnipotent God. He extended the thesis not only to the interaction of spirit and body (as to which Descartes may well have held a similar view), but to psychological causation generally. In this he seems only to have been more consistent than Descartes, since if God's continuous recreation of matter and its modes makes him the only true cause of corporeal change, it is difficult to see how finite spirits can be genuine agents any more than bodies are. For presumably spirits and their modes are no less subject to continuous recreation.

17

Another writer with whose somewhat tangled thoughts on these questions Locke must have been familiar (or, at least, whose thoughts are now difficult to disentangle) was Robert Boyle. Boyle agreed with Descartes that God conserved matter and upheld the laws of motion, but unlike Descartes he very much favoured the argument from design, holding it necessary to postulate that matter in law-governed motion was not 'left to itself, but skilfully guided at the beginning of the World' by the 'Divine Architect'. He rejected Descartes' view 'that chance should turn a chaos into a world', and there is no indication that he saw the laws of motion as flowing necessarily from God's essence. He did call them 'Signatures, and as it were Badges of [God's] Attributes', but the attributes he had in mind, power, wisdom and goodness, did not include immutability. Moreover, the principle that God acts in the simplest way was explained in terms of a crude and anthropocentric divine purpose:

His Divine Wisdom...[is] an attribute that may advantageously disclose Itself to us Men, by producing a vast Multitude of Things, from as few, and as simple, Principles, and in as Uniform a Way, as, with Congruity to His other Attributes, is possible. ¹⁸

Consequently there is no theological guarantee of the uniformity of nature:

And indeed, if we consider God as the author of the universe, and the free establisher of the laws of motion, whose general concurrence is necessary to the conservation and efficacy of every particular physical agent, we cannot but acknowledge, that, by with-holding his concurrence, or changing those laws of motion, which depend perfectly on his will, he may invalidate most, if not all the axioms and theorems of natural philosophy.¹⁹

The passage leads into the expression of an uncomplicated belief in the possibility of miracles, as events contrary to the laws of nature.

All this may suggest that for Boyle body is an entirely passive object of God's direct and continuous action, and that talk of the efficacy and powers of bodies is a metaphysically misleading way of talking of contingent 'laws of motion freely established and still maintained by God'. Yet in fact Boyle clearly did think that 'second causes' are really causes. He often described the universe as an 'automaton', while God was said to maintain by his concurrence 'those Powers, which he gave to the Parts of Matter, to transmit their Motion thus and thus to one another'. It seems that God's 'preserving concurrence' simply is his upholding bodies with the nature and powers which he gave to matter initially when he framed the laws of motion. That perhaps should not surprise us, since such traditional notions as *second cause* and God's *ordinary and general concurrence* had originally been invented for the sake of just such a fragile reconciliation of voluntarism with a commonsense conception of natural causality.²⁰ Yet the suggestion that powers have been arbitrarily and contingently stuck on to body is more than a little odd in the context of the critique of Aristotelianism. It is as if Boyle was inconsistently relying on his voluntarism to keep his ontology of matter clean, and on talk of an imposed 'nature' and 'powers' to avoid the notion that God is the sole and direct cause of every physical event. Yet in other remarks it seems possible to discern the influence of a rather different, more rationalistic model. No one, he wrote, 'has well performed the part of a true Physiologer, till he have circumstantially or particularly deduc'd the *Phaenomenon* he considers, by Intelligible Ways, from Intelligible Principles'. If we suppose such 'intelligibility' to extend to the laws of motion themselves, then it might seem that 'Nature acts...necessarily according to Laws Mechanical' in a sense which would encourage us to interpret God's freedom in his choice of laws as no different from his freedom to choose what substance to create 'at the beginning of Things, when there was no Substance but Himself, and consequently no Creature...by which He could be limited'.²¹

At least consonant with such a rationalistic line of thought, assigning an intrinsic intelligibility to the interactions of bodies, are Boyle's sceptical reminders of the boundaries of human knowledge. Why, he asked, does one motion in the brain produce a visual perception while another causes us to hear a sound? 'What can be answered, but that it was the good pleasure of the author of human nature to have it so?' This answer is simply 'the general one...that such is the nature of man'. Our being at such a loss shows that we lack 'a satisfactory account of the manner of sensation', just as the paradoxes of

infinite divisibility show that ‘we are yet to seek...the definition of a corporeal substance’.²²

It is not, then, easy to tell how much of what Boyle wrote represents the hard centre of his thinking and how much is fudge (or, indeed, whether such a distinction is appropriate). But two further points might be mentioned. First, he was explicit, like Descartes, that finite spirit is active: ‘God has left to the Will of Man the direction of many Local Motions in the Parts of his own Body...Man himself is vouchsaf’d a Power, to alter in several Cases, the usual Course of Things.’ Boyle himself drew an analogy between the mind’s interference with the motions of body and miracles, and he rejected the Cartesian suggestion that there would be something less quasi-miraculous about the supposition if we took human volition to modify the direction but not the quantity of motion involved.²³

The second point relates to gravity, which had been a difficulty for mechanist theories ever since ancient atomism, and which Boyle described as ‘perhaps that Quality, which of all others is most probably refer’d to an inbred Power and Propension’.²⁴ He himself generally assumed that things fall as they do because of some unknown mechanism, ‘either by the magnetical steams of the earth, or the pressure of some subtle matter incumbent on it, or by whatever else may be the cause of gravity’.²⁵ After Newton’s *Principia*, however, which ascribed to the inverse square law the same universality as to the laws of inertia and impact, Richard Bentley and others argued that that law is a special manifestation of God’s direct activity in the world.²⁶ Despite the strong tradition which assigned all the laws of motion to God’s arbitrary decree, the argument relied on an implicit contrast between the seeming contingency of the inverse square law and the intrinsic and manifest intelligibility of such laws as the laws of inertia. Newton himself maintained a sceptical neutrality on the question of whether gravity called for further mechanical explanation or should be assigned to God’s direct agency, although he was prepared to speculate in both directions. Yet in general the more a seventeenth-century philosopher was inclined to see gravity as a special case raising special problems (however to be solved), the more grounds we would seem to have to attribute to him a rationalistic presumption that *other* laws stand to reason in their own right.

The form of Locke's mechanism

The question of the form of Locke's own mechanism is more than a little complicated. A good place to start unravelling the issue is perhaps a set of passages concerned with the extent and causes of our ignorance of the natural world, passages which have been taken to constitute a direct assertion that at least some fundamental laws or connections in nature do not possess geometrical necessity and intelligibility, but are the brute effects of God's will. The bulk of these passages occurs in the context of the long argument of *Essay* IV.iii, 'Of the Extent of Human Knowledge'. First, in IV.iii.6, there is the famous claim that thinking matter is for us an open possibility: more precisely, that it is

impossible for us, by the contemplation of our own *Ideas*, without revelation, to discover, whether Omnipotency has not given to some Systems of Matter fitly disposed, a power to perceive and think, or else joined and fixed to Matter so disposed, a thinking immaterial Substance.

What this proposal boils down to may not immediately be clear, but some have taken Locke to have been asserting the absolute or ontological possibility that God should arbitrarily have made it the case that, whenever certain sorts of motions take place in the brain, conscious thought of a certain sort occurs; and that his setting up this mere correlation would count as his making the material 'system' the subject of the thinking, i.e. a thinking thing. Locke's claim that such a hypothesis is as intelligible as dualism would then rest on the point that dualism itself has to presuppose just such arbitrary and brute connections in order to account for sense-perception. That at any rate is what, according to this interpretation, Locke meant when he wrote:

Motion, according to the utmost reach of our *Ideas*, being able to produce nothing but Motion, so that when we allow it to produce pleasure or pain, or the *Idea*, of a Colour, or Sound, we are fain to quit our Reason, go beyond our *Ideas*, and attribute it wholly to the good Pleasure of our Maker. For since we must allow he has annexed Effects to Motion, which we can no way conceive Motion able to produce, what reason have we to conclude, that he could not order them as well to be produced in a Subject we cannot conceive capable of them, as well as in a Subject we cannot conceive the motion of Matter can any way operate upon? ²⁷

As the argument of IV.iii continued, Locke turned to the question of our knowledge of

specific substances or bodies such as flame and gold, i.e. our knowledge of what qualities coexist with the qualities which make up our complex idea of the substance in question. The reason why our universal knowledge is here 'very narrow and scarce any at all' is, first, because

the simple *Ideas* whereof our complex *Ideas* of substances are made up, are, for the most part, such, as carry with them, in their own Nature, no visible necessary connexion, or inconsistency with any other simple *Ideas*, whose *co-existence* with them we would inform our selves about.²⁸

But the reason, in turn, for this lack of 'visible necessary connexion' is two-fold. First, the observable qualities of a substance are all dependent on 'the primary Qualities of their minute and insensible parts; or if not upon them, upon something yet more remote from our Comprehension'. Since we do not know this 'root' from which the observable attributes grow, we cannot know which of them 'have a necessary union or inconsistency one with another'.²⁹ But a second, 'more incurable' factor

sets us more remote from a certain Knowledge of the *Co-existence*, or *Inco-existence* (if I may say so) of different *Ideas* in the same subject; and that is that there is no discoverable connection between any *secondary Quality*, and those *primary Qualities* that it depends on.³⁰

Here the contrast is with interaction at the level of primary qualities, including 'the change from rest to motion, upon impulse', which 'seem to us to have some *connexion* one with another'. By contrast, even if we knew the primary qualities of the 'invisible parts' we could not hope to 'establish certain and undoubted rules' involving secondary qualities. We 'can by no means conceive' how the former qualities can possibly produce in us the ideas of colours, tastes or sounds: 'there is no conceivable *connexion* betwixt the one and the other'.³¹

The same points are repeated later in the argument. It is agreed that,

if we could discover the Figure, Size, Texture, and Motion of the minute Constituent parts of any two Bodies, we should know without Trial several of their Operations one upon another, as we do now the Properties of a Square, or a Triangle.³²

The actions of certain acids on metals 'would be then, perhaps, no more difficult to know, than it is to a Smith to understand, why the turning of one Key will open a Lock, and not the turning of another'. Yet when we turn to secondary qualities the picture is different. That the primary qualities of bodies produce in us sensations of colour, sound etc., is 'evident'. But the 'mechanical affections' of bodies have 'no affinity at all' with the ideas of secondary qualities they produce in us. Our knowledge of such causality is limited to experience. We 'can reason no otherwise about them, than as effects produced by the

appointment of an infinitely Wise Agent, which perfectly surpass our Comprehensions'.
33

It is a fairly popular view that in these passages Locke was asserting the existence of brute, contingent, divinely ordained conjunctions. One recent commentator who holds it, Margaret Wilson, has recognized the existence of a conflict between what she here ascribes to Locke and things said elsewhere in the *Essay*. In particular she accepts that in his main discussions of substance and substances Locke adopted a geometrical model for the relation between real essence and properties. For example, if we could form an idea of the real essence of any sort of substance,

then the Properties we discover in that Body, would depend on that complex *Idea*, and be deducible from it, and their necessary connexion with it be known; as all the Properties of a Triangle depend on, and as far as they are discoverable, are deducible from the complex *Idea* of three Lines, including a Space.³⁴

In effect, Wilson accepts that Locke was, in such passages, a 'pure' mechanist. Such a strongly geometrical model seems, indeed, to have been presupposed in those sections in IV.iii itself, which explain why there is no 'visible necessary connection' between the observable qualities and powers of things, and how the discovery of the primary qualities of the minute parts might enable us to 'know without trial' how bodies will interact. Yet on Wilson's account Locke also asserted, in the course of the same argument, that some complex material bodies might have a power, the power of thought, which is *not* so related to their corpuscularian real essence, but is arbitrarily superadded to it by God.

How deeply such a speculation would seem to cut into Locke's anti-Aristotelian conception of the relation of attributes to the substance which possesses them appears from Leibniz's criticism of these passages in the *New Essays*. For Leibniz 'nothing unintelligible [ever] happens', and as for 'the good pleasure of our Maker...he produces and conserves in [bodies] only what...can be explained through their natures'. Leibniz continues

[If] we could not even conceive of a general explanation for the relations between soul and body, and if...God gave things *accidental powers which were not rooted in their natures* and were therefore out of reach of reason in general; that would be a back-door through which to re-admit 'over-occult qualities' which no mind can understand, along with inexplicable faculties...—helpful goblins which...do on demand anything...that a philosopher wants of them.³⁵

Leibniz's complaint was that the view which (anticipating Wilson!) he attributed to Locke not only introduces occult qualities as explanatory *dei ex machina*, but also misconceives the nature of the accidents or modes of things: it is 'to entertain a picture of little subsistent beings which can fly in and out like pigeons with a dove-cote. It is unwittingly to turn them into substances.' It is, moreover, 'to have recourse to miracles' of a kind which Leibniz evidently regarded as out of the question. It is no coincidence

that one of his criticisms of Descartes' attribution of force to God's agency had been that it postulates a standing miracle instead of accepting that force belongs to bodies themselves, constituting their innermost nature.

It is inherently implausible that Locke could have grafted the particular kind of voluntarism which Leibniz ascribed to him onto a conception of the relation between substance and mode which is based on the purest, most 'geometrical' form of mechanism. Where Wilson is satisfied to find such conflict even within one chapter and one argument, another commentator, Edwin McCann, has attempted to remove the contradiction by rejecting the interpretation of Locke as, even some of the time, a 'pure' mechanist. He proposes that the passages which seem to be setting out a strictly geometrical model for the deduction of powers from structure are in fact doing no more than drawing a broad analogy between geometry and mechanics: in each case a multiplicity of properties flow from a relatively simple nature, but the physical 'flowing' of qualities and powers from substructure is not *a priori* or geometrical. For McCann's Locke, we (or angels) could know how a body will behave 'without trial' even though there are not strictly deductive or *a priori* connections between a body's real essence and its observable qualities. What is necessary for such knowledge is only that there are contingent general connections between them, and that we know these connections from experience as well as knowing the real essence. In other words, McCann attributes to Locke something like the view which at any rate seems to be expressed in those passages of Boyle in which 'the laws of motion freely established and still maintained by God' are mentioned as if they were principles of explanation distinct from, and additional to, both the essential properties of body and its mechanical accidents.³⁶

The assumption that Locke might have been in agreement with Boyle on this point seems at first sight historically impeccable, and it can also be, at first sight, textually persuasive. The most promising passage is perhaps *Essay* IV.iii.29, or at any rate the first half of it:

In some of our *Ideas* there are certain Relations, Habitues, and Connexions, so visibly included in the Nature of the *Ideas* themselves, that we cannot conceive them separable from them, by any Power whatsoever. And in these only, we are capable of certain and universal Knowledge. Thus the *idea* of a right-lined Triangle necessarily carries with it an equality of its Angles to two right ones. Nor can we conceive this Relation, this connexion of these two *Ideas*, to be possibly mutable, or to depend on any arbitrary Power, which of choice made it thus, or could make it otherwise. But the coherence and continuity of the parts of Matter; the production of Sensation in us of Colours and Sounds, *etc.* by impulse and motion; nay, the original Rules and Communication of Motion being such, wherein we can discover no natural connexion with any *Ideas* we have, we cannot but ascribe them to the arbitrary Will and good Pleasure of the Wise Architect. I need not, I think, here mention the Resurrection of the dead, the future state of this Globe of Earth, and such other Things, which are by every one acknowledged to depend wholly on the Determination of a free Agent. The Things that, as far as our Observation reaches, we constantly find to

proceed regularly, we may conclude, do act by a Law set them; but yet by a Law, that we know not; whereby though Causes work steadily, and Effects constantly flow from them, yet their *Connexions* and *Dependencies* being not discoverable in our *Ideas*, we can have but an experimental Knowledge of them.

37

It is tempting to suppose that in this passage Locke first deliberately presented a contrast between geometrical relationships and dynamic, physical relationships, attributing the latter to God's providence as surely as any particular events. The conclusion which he might then seem to have drawn is that our knowledge of these contingently ordained laws must be experimental rather than *a priori*.

Such an impression of Locke's meaning cannot, however, be sustained. In this (one of, I think, only two places in the Essay where 'law' is used for the laws of physics) he was not asserting that our knowledge of fundamental laws is experimental, but that, given our ignorance of laws, our knowledge of observable physical relationships is experimental. The implications of this claim have been spelt out elsewhere. When there is no 'visible connection' between our ideas, so Locke said in IV.iii.14, there can be no universal knowledge: 'co-existence...cannot be perceived but either in particular Subjects, by the observation of our Senses, or in general, by the necessary *connexion* of the *Ideas* themselves'.³⁸ 'Experimental knowledge' of a law or of any universal generalization is therefore impossible in principle, a contradiction in terms. Experimental knowledge is restricted to particulars within our experience. All this of course ties in with Locke's firm distinction between knowledge and belief, and with his express doctrine that intuition and demonstration supply the only route to universal knowledge.³⁹

One source of difficulty in the interpretation of Boyle's remarks about laws and necessary connections is that they were not advanced in the context of an explicit and systematic epistemology; but there is no such problem in the case of Locke. McCann argues, against the 'geometrical' interpretation of Lockean necessary connections, that the locksmith who knows 'without trial' how a lock will respond to a key is surely not supposed to secure his knowledge by demonstration. But for Locke there was no way of securing *knowledge* in such a case which did not involve intuition or demonstration, and he was here appealing to the same thought as the authors of the Port Royal *Logic*: whether keys will turn locks or axles will turn mill-stones can simply be evident from the shape of the rigid parts, together with the idea of a body's defining capacity to push others out of its way. McCann is setting aside not only Locke's official epistemology, but also the tradition according to which mechanical change 'stands to reason' in its own right (granted that it can also be assigned to God's will at some level). To do the latter is simply to ignore the possibility that Locke, like Hobbes (if perhaps unlike Boyle), was a 'pure' mechanist.

Unlike Hobbes, however, Locke was not the dogmatic exponent of any particular mechanics. His preferred 'hypothesis' was a version of Boyle's corpuscularianism, which was not a determinate mechanics but a model which Boyle employed for largely destructive purposes against Aristotelian forms, occult powers and specious qualitative principles in alchemy. Sometimes, as in his chief discussions of primary and secondary

qualities, Locke endorsed that hypothesis virtually without qualification. Nevertheless it was in general overlaid, often emphatically, by the broad scepticism or agnosticism about essences, including the essence or 'substance' of matter, which figures so prominently in the chapter on our complex ideas of substances, *Essay* II.xxiii. Another chapter in which this agnosticism received special emphasis is, of course, IV.iii itself. Hence a major objection to our interpreting as strong voluntarism the latter chapter's presentation of the relation between mechanical processes and thought or sensation is that such a reading ignores the extent to which Locke found matter itself mysterious. To have claimed that a material 'system' could not possibly have the power of thought in the natural course of events (without, that is to say, a queer metaphysical act on God's part of the 'superaddition' of an independent or real accident) would have been to have made just the kind of dogmatic assertion about the natural possibilities which Locke was attacking in this very argument. In fact, 'superaddition' meant nothing so strange. In line with contemporary usage, Locke described entirely mechanical accidents, and in particular such mechanical 'perfections' as belong to animals and plants, as having been 'superadded' to matter by the divine architect.⁴⁰

Locke's position, then, was *not* that we know the nature of matter so well as to know that it could only think by a standing quasi-miracle. On the contrary, he held that, because we do not know the nature of matter or of thought, we do not know whether thought lies within the capacities of matter as 'fitly disposed' by God. The possibility of thinking matter lies open for us 'in respect of our Notions', but whether it is an ontological possibility is precisely what, according to Locke, we do not know. The theme of the argument of IV.iii.6, as of the whole chapter, is our ignorance of 'the Nature and inward Constitution of things', and that every substance 'has something in it, which manifestly baffles our understandings'. Just the same theme lies behind Locke's lamenting the lack of 'conceivable connexion' between primary and secondary qualities. His complaint was that the operations of bodies on thoughts, and of thoughts on bodies, are unintelligible to us because our ideas are not adequate to reality. We are convinced by experience that there is 'a constant and regular connexion' between them, but 'that connexion' is 'not discoverable in the *Ideas* themselves'. We are left to suppose that God 'has made them to be, and to operate as they do, in a way wholly above our weak understandings to conceive'.⁴¹ Even the presuppositions of the primary and secondary quality distinction are put in some doubt, for Locke allowed that the secondary qualities of things may not after all depend on the primary qualities of the minute parts, but on 'something yet more remote from our Comprehension'.⁴² There is nothing here to put in question Locke's agreement with Leibniz that God 'produces and conserves in [bodies] only what...can be explained through their natures'. His whole point was that we do not know those natures.

The section the first part of which seems to supply the strongest grounds for McCann's interpretation, IV.iii.29, is in fact full of the same hints of the epistemological contrast between our ideas of things and the things themselves. The problems over coherence, sensation and 'the original Rules and Communication of Motion' arise because 'we can discover no natural connexion with any *Ideas* we have'. We 'have but an experimental Knowledge' of causes and effects just because 'their *Connexions* and *Dependencies* [are] not discoverable in our *Ideas*'. Given that a primary message of the chapter concerns the

limitations on our stock of ideas, and consequently on our capacity to 'penetrate into the Nature, and inmost Constitution of things' by comparison with possible creatures better endowed than ourselves,⁴³ it would of course have been the grossest *non sequitur* on Locke's part to draw firm conclusions about the contingency of real relations from the lack of necessity in the corresponding ideal ones. The remainder of IV.iii.29 amply confirms that Locke's point was epistemological: 'From all which', it continues, "'tis easy to perceive, what a darkness we are involved in, how little 'tis of Being, and the things that are, that we are capable to know.' What we are incapable of is 'a philosophical *Knowledge*' or 'a perfect *Science* of natural Bodies (not to mention spiritual Beings)', and Locke uses the terms 'knowledge' and 'science' not only with emphasis, but with some precision. Demonstrative science is impossible because of our ignorance, not because there is nothing there to know.⁴⁴

There is, however, at least one respect in which the expression of the argument of IV.iii.29 is misleading. It opens, as we have seen, with a contrast between, on the one hand, 'visible' connections, exemplified by a geometrical relation, which no arbitrary power could make otherwise, and on the other hand those physical, causal relations which we cannot understand or explain through the ideas we have. Yet it would be a mistake to assume that Locke intended here to contrast mathematical knowledge as such with physical knowledge as such, or that the boundary between the two sorts of relation with which he is really concerned is anything but an epistemological one. For the same point as is made here is made in earlier formulations by means of physical examples on both sides of the dichotomy, rather than one physical and the other mathematical. In IV.iii.13 he managed the contrast between intelligible and unintelligible connections by granting a sort of dim appearance of necessity in mechanical interaction: 'the separation of the Parts of one Body, upon the intrusion of another; and the change from rest to motion, upon impulse; these, and the like, seem to us to have some *connexion* one with another'. Such changes are 'not beyond our Conception'.⁴⁵ The second formulation, in IV.iii.14, allows as an example of 'necessary dependence, and visible connexion' not only the uninteresting point that 'Figure necessarily supposes Extension', but also that 'receiving or communicating Motion by impulse supposes Solidity'.⁴⁶ Here Locke was endorsing the proposal in *Essay* II.iv, 'Of Solidity', that '*Impenetrability*, which is negative,...is, perhaps, more a consequence of *Solidity*, than *Solidity* it self.' As a primary quality solidity is an actual attribute, not a power, and the idea or sensation of solidity represents it as it is. It is 'the Idea most intimately connected with, and essential to Body', the attribute 'whereby we conceive it *to fill space*', and upon which 'depends [the] mutual Impulse, Resistance and Protrusion' of particular bodies.⁴⁷ All this may seem to take us very little way towards a determinate mechanics, but it is highly significant in so far as it demonstrates Locke's commitment to there being an intelligible and natural foundation for even the most fundamental powers (and *ipso facto* for fundamental laws) in the actual and perspicuous attributes of body itself. Thus the boundary really in question in the argument repeated in IV.iii.29 is not between mathematical and physical relations, but between those connections whose necessity is at least in part 'visible' to us, and observed coexistences which at best only probably point to a universal connection whose necessity is opaque to us.

McCann asks an interesting question: if Locke's point is simply epistemological, why did he claim, or seem to be claiming, that those physical relations which are opaque to us ought to be ascribed to 'the arbitrary Will and good Pleasure of the Wise Architect'? The ascription, McCann argues, is presented as a hypothesis to which we are and ought to be driven; but such a hypothesis would seem only to be in place if we could draw ontological conclusions about those cases in which dependable regularity is not explicable by 'visible' or 'conceivable' connections. Why else, it might be asked, should we be enjoined to treat unintelligible physical regularities as on a par with acknowledged contingencies, and as being as dependent on God as the resurrection of the dead or the ultimate fate of the world?

Fortunately, there seem to be two available answers to this rhetorical question. The first relates to the tactics of Locke's argument. In the discussion of thinking matter he argued in effect that where there is no contradiction in our idea of a thing, there we ought to assume that that thing is in God's power. To assume that, just because we cannot understand how matter should think, thinking matter is impossible, is to argue in a way which 'becomes not the Modesty of Philosophy'.⁴⁸ To Stillingfleet Locke used stronger language: to accept the immaterialist argument is not only 'to set bounds to God's omnipotency' on the basis of our 'narrow conceptions', but commits us to the absurdity of denying the occurrence of any event which we do not understand.⁴⁹ Hence for Locke whatever is not 'visibly' impossible or self-contradictory ought piously to be regarded as within God's power to create—even though, for all we know, it might be in itself impossible. The obverse of this principle of piety is that whatever is not visibly necessary ought to be presumed *not* to be one of those immutable relations or connections which even God could not sever. On this interpretation, then, Locke was saying that, despite the epistemic possibility that the connection is necessary, it ought in piety to be regarded as dependent on God's will. Locke was not telling the reader to adopt a certain hypothesis for its own sake, but enjoining Christian humility in matters of unproven impossibility and necessity.⁵⁰

The second possible explanation of Locke's motive would take him less to be recommending us to ascribe experienced regularities to God's will than to be giving an account of people's actual practice. Boyle had made the point, as we have seen, that what people do not understand, they ascribe to the good pleasure of the Author of the Universe: 'when we give such general Answers, we pretend not to give the particular Physical Reasons of the things propos'd, but do in effect confess we do not know them'.⁵¹ Boyle seems to have failed to apply the point to his own ascription of the laws of motion to the will of God, but there was nothing to stop Locke from extending it to fundamental laws as his own critical comment on Cartesian voluntarism, and conceivably even on Boyle himself.

Which of these two answers to McCann's question is preferred may depend on how far we read Locke as *enjoining* us to ascribe those coexistences which are not understood to God's will, rather than as simply *describing* what people in general, and certain philosophers in particular, are inclined to do anyway. It is possible that Locke had both intentions: the attitude natural to us in ignorance of necessary connections is also the properly pious one. But whichever is his motive, the spirit in which he allows that 'we

cannot but ascribe' the laws of motion and impact 'to the arbitrary Will and good Pleasure of the Wise Architect' is a very long way from the spirit of Descartes' dogmatic voluntarism, or of Bentley's proof of God's existence from the supposedly known contingency of the law of gravity. On Descartes' view, we can understand how specific laws of motion flow necessarily from God's nature and power. For Locke, if we ascribe such regularities as we observe to God's direct agency, we do so only because we at best dimly glimpse how they might flow necessarily from the dimly understood nature of the matter which God has arbitrarily created.

There might seem to be evidence in a passage from *Some Thoughts concerning Education* that Locke accepted Bentley's estimation of the significance of gravity. Here he remarked with regret that physics can encourage atheistic materialism,

when yet it is evident, that by mere Matter and Motion, none of the great Phaenomena of Nature can be resolved: to instance but in that common one of Gravity, which I think impossible to be explained by any natural Operation of Matter, or any other Law of Motion, but the positive Will of a Superiour Being, so ordering it.⁵²

The question arises, however, whether Locke here saw the 'positive will' of God as engaged in a continuous and quasi-miraculous or supernatural interference with the 'natural operations' of matter, or as having so ordered the universe that, given those natural operations, the phenomenon of gravity ensues. Since he went on to speculate that the Flood might have been due to 'God's altering the Centre of gravity in the Earth for a time (a thing as intelligible as gravity it self, which, perhaps a little variation of Causes unknown to us would produce)', it seems more than possible that Locke had in mind the familiar figure of the divine architect rather than the perpetual miracle-worker. Certainly, other passages which reflect his response to Newtonian gravity carry no commitment to a notion like Bentley's of God's immediate agency, but rather convey a strictly epistemological, sceptical message. In the fourth edition of the *Essay* he replaced his earlier claim that the way bodies operate on one another 'is manifestly by impulse, and nothing else', since they 'cannot operate at a distance', with the weaker assertion that impulse is 'the only way which we can conceive Bodies operate in', i.e. which is intelligible to us.⁵³ His comment to Stillingfleet was similarly epistemological: Newton's 'incomparable book' had convinced him that 'God can, if he pleases, put into bodies powers and ways of operation above what can be derived from our idea of body, or can be explained by what we know of matter.' God has arranged gravity 'by ways inconceivable to me'.⁵⁴ Finally, no ontological conclusions are drawn when the contrast between the intelligible, axiomatic laws of inertia and the seemingly contingent inverse-square law is set out in what might be supposed Locke's most considered statement, the posthumously published *Elements of Natural Philosophy*:

it appears, as far as human observation reaches, to be a settled law of nature, that all bodies have a tendency, attraction, or gravitation towards one another....

Two bodies at a distance will put one another into motion by the force of

attraction; which is inexplicable by us, though made evident to us by experience, and so to be taken as a principle in natural philosophy. ⁵⁵

If the account here offered is correct, then Locke's mechanism was 'pure', but overlaid with doubt. Corpuscular mechanism was for Locke the best inadequate hypothesis about bodies we have: he feared

the Weakness of humane Understanding is scarce able to substitute another which will afford us a fuller and clearer discovery of the necessary Connexion, and *Co-existence*, of the Powers, which are to be observed united in several sorts of them. ⁵⁶

Observable mechanisms provide a kind of glimmering of what the ideal mechanical science would look like, but no more. Consequently Locke's doctrine does not in fact satisfy the provisional definition of 'philosophical mechanism' offered at the beginning of the previous chapter, but we may take that to reveal a shortcoming of that definition. If we can be sure that Locke is a 'pure' mechanist, we must add that his mechanism was a formal commitment, not a material commitment to any existing geometrical mechanics. He might be described, then, as a 'pure ideal' mechanist.

That Locke's God does not continuously and directly determine the motions of bodies (not even by a 'general and ordinary concurrence') does not mean that Locke was in no sense a 'voluntarist'. His God chose the laws of mechanics (whatever they are) in choosing to create matter (whatever that is), and that divine choice can be seen from at least one point of view as *more* arbitrary and contingent than the decrees of the Cartesian God, which have flowed eternally and necessarily (if 'freely') from the divine immutability. For both philosophers, it makes sense for us to suppose that matter could just as well have obeyed such and such a set of laws as such another set. According to Descartes' voluntarism, that is because either is equally possible ontologically in relation to the known geometrical essence of matter; although only one set of laws, which Descartes purported to identify, is compatible with the essence of God. According to Locke's voluntarism, on the other hand, the possibility is epistemic. It exists because we do not know the essence of the substance which God has arbitrarily chosen to create (and which we call 'matter'); although only one set of laws, also unknown to us, flows necessarily from the unknown essence of that substance. We shall find both an important similarity and an important difference in his treatment of the moral 'Law of Nature'.

Reflections on rationalism, empiricism and mechanism

'Rationalism', 'empiricism' and 'mechanism': how are these terms related? First, all forms of philosophical mechanism, 'pure' and 'impure', are forms of rationalism. The term 'mechanism' has perhaps been sufficiently explained above. A good definition of 'rationalism' might be expected to account for the notorious difficulty there is in classifying philosophers neatly as 'rationalists' or 'empiricists'. In what follows the term will be used in the first instance for the view that the world is an intrinsically intelligible place. Rationalism in this sense assumes that the occurrence of change is in accordance with fundamental rules which are at some level necessary; and that their necessity is of a kind which can in principle be grasped *a priori* by the mind, if not by finite human minds. According to a particularly strong form of rationalism which will not be discussed here, namely Spinoza's, even existential propositions are (if true) ultimately necessary and in principle intelligible, if not to us. Here the expression 'strong rationalism' will be used for a different view: the epistemological view that some appropriate method enables the human intellect to arrive at conscious knowledge of the necessary laws of natural change.

It has been said that for rationalism these laws or fundamental principles of change are 'at some level' necessary because, as we have seen, an 'impure' mechanist in the Cartesian mould need not regard laws as intrinsically necessary, but only as necessary in relation to the nature of God. So an impure mechanist could be a strong rationalist. Conversely Locke, as a 'pure ideal' mechanist, was a weak rationalist: 'weak' because, although he strongly held to a rationalist ontology, he believed that for the most part the necessity intrinsic to all change cannot be known or understood by us. There can be different kinds and degrees of 'weakness', as of 'impurity'. Locke, as we have seen, believed that we can achieve some fraction of an understanding of necessities, and his sceptical arguments do not rule it out absolutely that some speculative genius should crack the code, achieving angelic or God-like grasp of natural change.

Not all rationalists, on the present definition, were mechanists. Perhaps Aristotelians could be counted as a sort of non-mechanistic rationalists, but Leibniz is a clearer example. He held, as we have seen, that all change flows intelligibly from essences, but argued that essences must in the end be conceived of dynamically, as forces for change, rather than in terms of purely actual, non-dynamic properties, as mechanism conceived of them.

The distinction between 'rationalism' and 'empiricism' is not a sharp one not only because the same philosopher may be both a rationalist and a concept-empiricist, as

Locke was, but for two reasons more pertinent to the present argument: first, because empiricism, like rationalism, has ontological and epistemological faces which can come apart; and second, because ‘impure’ mechanism is possible, founded on more or less strong voluntarism. Both of these reasons need exploration and illustration. The supposedly empiricist view that natural connections between events are mere brute regularities, unintelligible in principle, has generally been credited to Hume; yet if what is meant by ‘unintelligible’ is *intrinsically unintelligible in and by themselves*, then it will be seen that impure mechanism, although it might be a form of strong rationalism, already embodied such a view. Hume drew directly not only on Berkeley’s extreme, empiricist voluntarism, but also on the antecedent Cartesian tradition (especially Malebranche’s occasionalism) in so far as it proclaimed the intrinsic inertness of matter. Where, clearly, he differed from that tradition was in his rejection of the role assigned to God’s will: in particular, his rejection of the view that the laws of mechanical change can become intelligible to us through reflection on God’s immutability, or by recourse to such teleological rules as the principle that God produces the greatest possible variety by the simplest possible ways. Even so, Hume cannot be set down as, so to speak, a voluntarist without God. For he did not deny (but, rather, asserted or implied) that natural change is intrinsically necessary. He simply held that the necessity is utterly beyond our conception. In the world as we can experience and conceive of it, there are, he held, only brute regularities. But among things as they are in themselves, he assumed, there is an intelligible order.

Hume was, as one perceptive commentator has put it, a sceptical realist,⁵⁷ but his realism has commonly been overlooked. One reason for this is that he was unlike Locke, whom in some other important respects he resembled, in allowing the ideas gleaned from experience no possible value as materials for speculation about the world beyond experience. Our ideas are incapable of affording the least glimmerings of necessity, although experience of regularity gives rise to a sort of surrogate and confused idea of necessary connection. That alleged limitation on our thought causes difficulty when it is suggested (as the sort of sceptic that Hume was must suggest) that there exists an intelligible order of which we have no knowledge. If we have no conception of it, how can the sceptic himself conceive of its existence? Later empiricists, such as Russell at least some of the time, took the tough line that the world as we experience it is all there is. Their view might be called ‘strong’ empiricism and Hume’s (despite potential scandal) ‘weak’ empiricism (although it was not as weak as Locke’s).

A weak empiricist, it will be seen, can also be a weak rationalist, since weak empiricism is epistemological empiricism, while weak rationalism is ontological rationalism. Strong (or epistemological) rationalism implies weak (or ontological) empiricism, and strong (or ontological) empiricism implies weak (or epistemological) rationalism, but in neither case does the reverse implication hold unless with the aid of some further principle. Hence the epithets ‘strong’ and ‘weak’. All this helps to explain the desperate inadequacy of any simple dichotomy between ‘rationalists’ and ‘empiricists’.

These preliminaries completed, we can turn to the task of assessing Locke’s mechanism as a form of rationalism. A reasonable start might be further comparison of

his theory with Hume's. It might also be reasonable to start with the similarities and differences between their forms of rationalism as such, before moving on to the more specific issues raised by pure mechanism and Hume's criticisms of it.

Although Hume, like Locke, was an ontological rationalist, his epistemological empiricism is so rigorous that the role of rationalism in his argument is limited in effect to philosophical asides which are necessary to the structure of his sceptical realism but which, according to his own theory, should be held illicit or, indeed, meaningless. His ontological rationalism, in other words, plays no role in his account of the causality of our ordinary or scientific beliefs. It is thus about as weak as any form of rationalism could be. He did not deny that we have a belief in the uniformity of nature, or that every event has a cause, but he ascribed this belief, like all our causal beliefs, to custom or habit, an irrational principle of the imagination. He took it that its contradiction makes perfect sense to us. For Locke, however, the principle of universal causation is evidently necessary in itself, and in harmony with all our reasoning about the world. Its necessity is no different from the necessity that all changes flow from the essences of things, or that behind every observed change lies 'a Power somewhere, able to make that Change'.⁵⁸

Among later empiricists it was perhaps Russell who most effectively seized on this point of difference in the course of developing a revised version of Hume's view. The revision took the form of rejecting Hume's sceptical ascription of causal reasoning to habit, and reconstituting his account as an account of the conditions of rational inference from the observed to the unobserved. The supposedly self-evident principle of such inference he called the 'principle of induction'; which, as it applies to laws, he set out as follows:

(a) the greater the number of cases in which a thing of the sort A has been found associated with a thing of the sort B, the more probable it is (if no cases of failure of association are known) that A is always associated with B;

(b) Under the same circumstances, a sufficient number of cases of the association of A with B will make it nearly certain that A is always associated with B, and will make this general law approach certainty without limit.⁵⁹

According to Russell, our belief in the uniformity of nature, although a 'general principle of science', is subordinate to our acceptance of the principle of induction. It is at best probable, being itself based on experience and inductive reasoning.

This difference between Locke's traditional view on the one hand and Hume's and Russell's views on the other constitutes one of the great watersheds of epistemology, and assessment of Locke might well begin here. There seems little doubt that his approach in the end proves preferable to that of his critics.

One familiar set of problems arising out of the subordination of the principle of uniformity to a supposed principle of induction surrounds the question of how the former could possibly be either supported or refuted by experience. If what counts as a confirming instance of the principle of uniformity is a success in explaining an event by bringing it under a law, and if no law is more than probable, then it can only ever at best be probable that we have a confirming instance. More seriously, what would constitute a

counter-instance? Not just our failure to explain an event, for then the truth of the principle would depend on our skill and luck as scientists. The failure would need to be attributable to reality rather than to ourselves; but to do that, of course, would beg the question.

More seriously still, Russell palpably failed to state an acceptable 'principle of induction', taken to be a principle of inference, and no one since has succeeded where he failed. A number of well-known paradoxes follow from his formulation. For example, if a generalization subtends a range of subordinate generalizations, it is possible that all the observed instances of the wider generalization should also be instances of one particular subordinate generalization. According to Russell's principle, the two generalizations possess the same probability, yet it is clear that the wider one incurs the greater risk. Second, a generalization cannot be less probable than a wider generalization which logically entails it, so that, despite what Russell says, some generalizations of which there are no observed instances at all may be as probable as others of which we have observed many instances. Third, a difficulty for the principle arises from the case in which one observer has observed the same number of instances of a generalization as another observer has observed of a narrower generalization falling under it. Both observers may have observed the same number of instances of the wider generalization, yet if the instances observed by the first observer are in some sense properly representative (e.g. if observations in conformity with Boyle's Law are made of a wide variety of gases in widely variant circumstances) then it is evident that the first observer may have more reason than the second observer to accept the wider generalization. Consequently he will have more reason to accept the narrower generalization too, since it is deducible from the wider one. Fourth, and perhaps most spectacularly and notoriously of all, 'All *As* are *Bs*' is equivalent to its contrapositive, 'All non-*Bs* are non-*As*', so that whatever supports the one supports the other. Thus Boyle's Law, according to Russell's principle, would be confirmed by the observation of anything which neither conforms to Boyle's Law nor is a gas, such as a lump of lead. This consequence, found acceptable by some philosophers, seems plainly incredible.⁶⁰

In the face of these and other criticisms modern empiricists have attempted to patch up the 'principle of induction'. For example, the breadth of the range of observed instances has been added to their number as a determinant of the degree to which experience confirms a hypothesis. Another determinant proposed has been the simplicity of the hypothesis in relation to rival hypotheses. The introduction of such independent variables removes the attractive rigour of Russell's account, but also raises the problem of the unity of the principle. We cannot accept a mere hotch-potch of unconnected intuitions as a principle of rational inference. Sometimes this requirement has been recognized, as in the confused attempt to reduce the consideration of range to that of the number of instances observed by treating subordinate generalizations as instances of wider ones.⁶¹

In general, however, empiricists have optimistically assumed that Hume and Russell were at bottom correct, an assumption founded, perhaps, on the following thought. Since we recognize a distinction between rational and irrational expectation (or belief in a generalization), and since we can commonly agree in our application of the distinction in particular cases, including the cases brought as paradoxes for Russell, it must be possible,

given sufficient analytical skill, to state the criteria we use: when stated, they will constitute a principle of rationality, no doubt built on Russell's, which can be accepted as an independent principle of induction.

Such optimism is misguided. For ontological rationalism does not of course imply that it is impossible to say anything about our criteria (or the proper criteria), whether in general or in particular sorts of circumstances, for treating a particular hypothesis as well-confirmed. It holds rather that these criteria become intelligible as reasons for belief only against the presumption of a *natura rerum*, a law-governed reality some aspect of which the hypothesis is intended to capture. Science is explicit speculation about the natural order, but all our reasoning about the world, however informal, assumes the existence of that order. Such a view can explain why different considerations carry weight, and different procedures are called for, at different stages of inquiry, or depending on the subject-matter or special conditions of inquiry. Some procedures for assessing evidence are mathematical, but the ideal of a fundamental, universally applicable mathematical logic of confirmation is an *ignis fatuus*. Sometimes an estimate of probability may be rationally determined by the number of cases observed, but that is in very peculiar circumstances, such as the tossing of an unexamined die or coin, in which the crucial evidence intelligibly lies in that consideration. Equally, there are rational procedures for gathering and assessing evidence in order to identify significant correlations, but the appropriateness of these statistical procedures is again dependent on the level of inquiry: they do not constitute a universal scientific method. Another level or stage of inquiry may call for just one carefully constructed experiment or set of experiments. Sometimes the result of such an experiment will be no more than suggestive, encouraging or discouraging a line of investigation. Sometimes a fully determinate hypothesis will be refuted, or else given something like the backing Descartes envisaged in his analogy with the speculative interpretation of a code which converts a jumble of letters into the Lord's Prayer.⁶²

Paradoxically enough, empiricism is notoriously weak in its philosophy of experiment. If the rationality of our expectations and general beliefs derives entirely from past experience, why should we bother to be selective about the experiences that we are about to have? For all that Russell says, we might as rationally attempt to preserve the probability of a generalization by preventing the novel circumstances most liable to produce a counter-instance, as test the generalization by deliberately bringing just those circumstances about.⁶³

This failure of empiricism is associated with another equally notorious: a failure to distinguish between what actually occurs and what necessarily occurs; or (in a related way) between what actually happens and what would happen in certain circumstances. As long as the universal propositions of science are thought of as mere summaries of what actually occurs, the point of experiment remains philosophically opaque. The notion of natural law, on the other hand, and those of natural necessity, possibility and impossibility, are notions of relationships which are not contingent on circumstances or on anything that human beings might choose to do. Such relationships are there to be *revealed* by human activity, by experiment and directed observation. The significance of the activity, and the grounds of the distinction between its being well or badly done, can

only be understood on the presumption of a fundamentally unchanging nature, i.e. on the presumption that each experiment is of universal significance and that 'unrepeatable' results are so only in so far as experiments have been inadequately described.

The difference between 'rationalism' and 'empiricism' here in question could be summed up as follows. Consider the inference, of a kind we make all the time both in science and out of it, from the premise that two states of affairs are, as far as can be observed, qualitatively identical to the conclusion that they will probably have the same outcome; or, conversely, from the premise that two states of affairs have had observably different outcomes, to the conclusion that there must have been some difference, unobservable if not observable, between them. Such inferences stand or fall with the principle of the uniformity of nature. For rationalism, since the proposition that the world is not law-governed does not make sense, they are intrinsically rational, requiring no justification outside themselves. For empiricism, they are justified only in the context of an inductive justification of our belief in the uniformity of nature. Of the two, rationalism would appear to have incomparably the better case, for it does indeed seem intrinsically reasonable to assume that in like circumstances like events are to be expected; and, for the reasons sketched out above, extremely difficult to imagine how that assumption could be 'inductively' based on experience. On the contrary, it seems evident that the principle was as rationally presupposed in the earliest days of science or human thought as it is now.

It might be objected that the principle of the uniformity of nature, so far from being ineluctable, is probably false, since quantum mechanics supplies the fundamental theory best supported by induction. According to quantum theory, there is only ever a slender probability that two identical states of a closed system should have identical outcomes. Do we not therefore have good inductive reason for rejecting the principle of uniformity? To this it may be replied that quantum theory rests on formally defined 'inductive' reasoning no more than any other theory. Like other theories, it is a hypothesis about the rationally assumed *rerum natura*, accepted in so far as it seems the best explanation of the phenomena. Just because quantum laws are probabilistic, there is some difficulty for the belief that they are ultimate. Yet, even if physicists adopt that belief, their reasons for accepting quantum theory rely, in effect, on the assumption that nature is law-governed, that things all and always behave in accordance with their natures, and that experimental results are acceptable only in so far as they are repeatable. The view that quantum laws are ultimate is paradoxical, not because it casts doubt on that assumption, but because it is at least *prima, facie* difficult to make sense of the claim that there are irreducible objective probabilities.

Rationalism, it is true, suffers from a notorious difficulty. In denying that the principle of uniformity is contingent and empirically grounded, it is called upon to explain how a proposition seemingly about the world can be known without sensory recourse to the world. The doctrine that the knowledge is innate or in some other way God-given is an answer that was as useless to Locke as it is to us. On the other hand empiricism faces a comparable problem with respect to the status of the 'principle of induction'. In virtue of the linguistic theory of logical necessity many modern empiricists have been committed to the view that the truth of the principle is a matter of the meanings of words, so that the

principle is in effect a definition of the word 'probable' and its synonyms. Yet Russell's principle and every proposed amendment are evidently normative, intended to capture what is reasonable rather than what the evaluative word 'reasonable' means. If there were a stable and tenable principle of induction, it would be no less a standing counter-example to the linguistic theory of the *a priori*, one of the main tenets of twentieth-century empiricism, than an *a priori* principle of uniformity would be.

Each of these problems can lead on to a would-be solution on idealist or pragmatist lines. Kant, of course, saw the principle of the uniformity of nature as a synthetic *a priori* truth, a principle true of the world as we experience it just because it reflects the contribution of the mind to the construction of that world, rather than the contribution of things as they are in themselves. Present-day neo-idealists have modified Kant's theory of the synthetic *a priori* in three main respects. First, such principles are regarded as essentially mutable, or open to replacement; second, in so far as they are regarded as structural, they are taken to reflect the structure of our language, 'theory', or system of beliefs rather than the structure of the human mind; third, it is characteristically held, following Quine, that distinctions cannot be drawn between necessary and contingent, analytical and synthetic, or *a priori* and *a posteriori* propositions except in so far as there are some propositions central to our belief-system which we are less inclined to give up than those at the periphery. The foundation of a notion of a synthetic *a priori* thereby evaporates, although it may be agreed that an acceptance of the principle of the uniformity of nature helps to constitute what we currently count as rational. As such, the principle may have a pragmatic justification. Like the rest of our 'theory', it is there because it has worked, although it might be replaced by something which works better. Such loose and easy philosophical explanation leaves it unresolved, however, whether the principle of uniformity should be regarded as basic, rather than posterior to some other principle. For just the same pragmatist pater can be given by those inclined to follow empiricism in according priority to a 'principle of induction'.

Current pragmatist orthodoxies are criticized at what is perhaps sufficient length elsewhere in the present work. But Kant's claim that we bring the principle of uniformity to experience, rather than derive it *from* experience, is insight into more than a contingent feature of our mutable 'conceptual scheme'. The principle is *a priori*, and is synthetic in that it is not analytic, but it is not in any ordinary sense 'about the world'. It is like the principle that time-travel is impossible, or that nothing is larger than itself. To suppose otherwise, to suppose exceptions, fails in the last analysis to make sense. A strong argument here is Kant's own 'Refutation of Idealism':⁶⁴ there is a mutually essential relation between subjection to natural law and objective existence in space and time. It is no accident that the best clocks are constructed in the light of our knowledge of fundamental physical laws.

One route towards further uncovering the necessity of the Principle of Uniformity might begin with the consideration that every law must be connected with every other law. It is perhaps clear enough that no world could contain, for example, just the laws of classical mechanics and certain astrological laws, quite unrelated to mechanics, which determine that certain individuals will be lucky, others fall into danger and so forth. For let us suppose such a world, and that someone born under a lucky star in that world

stumbles over a cliff and is brought by his star to rest in a projecting bush. Now we have to make sense of a division of labour between mechanical and astrological causation such that the astrological influence does not produce a counter-instance to the laws of mechanics, is not itself mechanical, and is not an instance of a wider law which in some unimagined way incorporates both mechanical and astrological laws. Yet this cannot be done for the same reason that we cannot make sense of the proposition both that Newtonian physics is true without qualification and that (say) discrepancies in the perihelion of Mercury are to be explained by some law having no relation to the laws of Newtonian physics. Another parallel is the difficulty facing the dualist who proposes that, while most physical events have physical causes, some have psychological causes instead. A physical law with such quasi-miraculous exceptions is no law at all.

It might be thought that there could be unconnected laws which governed changes different in kind, as motion is different in kind from change of colour. That would be to ignore the objectivity of sense-relative attributes. The colour of an object cannot just be its visual appearance, which must have some determinate physical basis in both the object and the perceiver. To suppose that the colour of an object is causally or nomically connected with none of its other attributes, as if a change in colour could occur without any other change, is to suppose away what makes the colour an attribute of that object, and what distinguishes it from mere illusion. We need to give content even to the postulate that both sorts of change take place in the same world: which no doubt the empiricist will do by taking it that both might be perceived by the same observer. Yet this move itself presupposes physical mechanisms of perception and of unitary consciousness. Membership of the same world means subjection to the same system of law. An object's or quality's bare existence in space is not separable from its reciprocal causal relations with other objects, including observers. If we postulate the spatial existence without the causal relations, we suppose that it makes no difference where the object is. In effect, we have supposed away the basis for assigning the object to one place rather than to another, or to any place at all.

Argument of this general type, which includes the Kantian appeal to the preconditions of determinate time, demonstrates the necessity that the world is law-governed. Yet it may nevertheless seem to leave it open whether the laws themselves are necessary or contingent. We have seen that rationalists of a strongly voluntarist persuasion could hold that physical laws are not intrinsically necessary, but are intelligible only in relation to the nature of God. Such rationalism was put to the test by the issue of miracles, for the more it was stressed that laws are arbitrary and contingent, the more natural was the theologically orthodox notion that they might be suspended. More than one philosopher appears to have been torn between his rationalism and such easy orthodoxy, or to have been trying in effect to accommodate both. The forms of voluntarism and mechanism which have here been ascribed to Locke, however, would make it absolutely impossible for God to create matter either not subject to law in general, or not subject to the specific laws (whatever they may be) to which it is in fact subject. For both necessities spring, on his view, not from the nature of God, but from the nature of what God has arbitrarily created.

It may perhaps be enlightening to consider how much in each of these positions

depends on the role of God. For 'impure' mechanism, to remove God would be to remove not only the ground of the intelligibility of the laws of nature but the ground of their existence as his general volitions. Locke, on the other hand, placed both the intelligibility and the existence of laws in things themselves, in mechanical or quasi-mechanical essences. His view might well appear to have been the better equipped to survive the eclipse of theological voluntarism, yet, oddly enough, the doctrine that laws are in themselves contingent has dominated much modern philosophy of science. Natural corollaries of that doctrine, once God has been excluded, are the denial that laws are at any level 'intelligible' and the conception of laws as having no existence apart from their instances. A law-statement, on the neo-Humean view (if that adjective is not unfair to Hume), is a blank summary-cum-prediction which 'explains' only in so far as it gathers up occasions for the purpose of inductive inference. So it is after all not easy to combine the view that there is necessarily a comprehensive system of law with the view that specific laws are contingent. Once the notion has been adopted that the terms of a law, the items nomically connected, are thin, natureless events or states of affairs logically and ontologically independent of the law and of each other, arguments for the necessity of the uniformity of nature are likely to appear sophistical. For how could it be necessary that independent items should be nomically ordered? If nomic order is a precondition of determinate position in space and time, then that, it might seem, is so much the worse for 'space' and 'time', which must therefore be contingent constructs built out of a contingent order. To approach the same point from another direction, the argument against inductivism requires completion by an argument against the inductivist's ontology. Things in the world are not to be captured, like Berkeley's 'inert' ideas or Hume's independent impressions or Russell's sense-data, by the names of sensible qualities alone, but by a linguistic network the parts of which are essentially related to one another.

If this last remark sounds like idealism, that is because idealism has presented itself, in one form or another, as the only alternative to empiricism. In its fundamental features the linguistic network we employ is not something imposed on the world by 'the mind' or by us, as idealism has it, but is generated by the condition of our existence in the world as beings with cognitive faculties. As it has been argued above with respect to the primacy of substance (an argument which was begun in Volume I and will be completed in the discussion of identity), the fundamental structure of language, determining what at the most basic level it makes sense to say, is in no way arbitrary, but derives from the independent existence and identity of the logically basic objects of our knowledge. We do manage cognitively to hook on to things which are independent of our thought and experience of them. Our language is necessarily built upon that achievement, although language also reflects its being through sense-experience that those independent realities are captured as objects of thought under the category of substance. That is why a world of natureless, inert Humean impressions does not make sense. It is not just because the mind gives itself no time to focus on blank unconnected data before hurrying to impose on them some inductively useful 'theory'.

If, then, we must, with Locke, ground physical agency and law in the natures of what is subject to law (understanding law in terms of essence, rather than essence in terms of

law) the question arises how the necessity of laws is to be regarded, i.e. the necessity by which things, being what they are, *must* interact as they do. In particular, is nomic necessity the same in kind as geometrical or logical necessity? For Locke, as we have seen, there is a difference between two kinds of case corresponding to the difference between *natural history* and *science*, between nominal essence and real essence or, as we might say, between experience and theory. Universal propositions at the first level, e.g. the proposition that gold is soluble in *aqua regia*, are for him contingent, unless true by arbitrary definition. Even on a different and better account of the meaning of 'gold' than Locke's, an account according to which it is an indefinable general name of a natural kind, something like that conclusion might seem still to stand. Since we can identify and name gold pre-theoretically without knowing its nature, it follows that, even if it is in some sense necessary that gold is soluble in *aqua regia*, that is not something anyone can know *a priori* in knowing the meaning of 'gold'.⁶⁵ In Volume I it was suggested that *a priori* knowledge and the notions of 'logical' necessity, possibility and impossibility are best understood as concerned with 'what makes sense', and to presuppose representation, paradigmatically, but not necessarily, linguistic representation. For example, we look at a picture of an Escher staircase, or read a description of one, and recognize that the representation does not make sense, i.e. its object is impossible. Yet it would appear to make perfectly good sense to suppose, however falsely, that gold does not dissolve in *aqua regia*. The necessity with which it does so dissolve, the impossibility that it should not so dissolve, would seem to have nothing to do with representation. It would therefore seem that natural necessity should be sharply distinguished from logical necessity.

There is, however, an argument on the other side, for the conclusion that 'Gold dissolves in *aqua regia*' is indeed logically necessary. For the meaning or linguistic role of 'gold' is to be the name of gold (this stuff) as it actually is. We can suppose a logically possible world in which the name 'gold' is employed for some substance so structured as not to be soluble in *aqua regia* (as *aqua regia* actually is), but then 'gold' as used in that world would not mean the same as 'gold' as used in the actual world: it would not be the name of gold. So it would seem logically impossible that *gold* should not dissolve in *aqua regia*, even if that cannot be known *a priori*. The logical impossibility is not due to some formal or implicit nominal definition of 'gold' as (say) *yellow, heavy metal, soluble in aqua regia*, but to the semantic connection between 'gold' and gold, a stuff which is of its nature soluble in *aqua regia*.

Before attempting to adjudicate between these two arguments, it may be helpful to consider the other kind of case, involving so-called 'theoretical' entities. These (supposing the theory in question to be true) are not less real than observable things and stuffs, but they are identified in a different way. Electrons, protons, neutrinos and the like are identified only through a speculative theory embodying the principles of their behaviour and effects. Many neo-idealists or pragmatists take identification through a scientific theory as their paradigm of identification in general, in keeping with their doctrine that all language is theoretical. In this they are wrong: it is not necessary to have a theory about gold in order to identify and refer to gold. But it follows that there is an even stronger case than with respect to 'all gold is soluble in *aqua regia*' for taking law-statements involving theoretical entities to be 'logically' or 'conceptually' necessary, and

indeed to be knowable *a priori*. Very roughly, a theory is implicit in the notion of an electron and, in so far as electrons cannot be picked out or identified independently of the theory which defines what an electron is, an assertion of the existence of electrons stands or falls with the theory. Thus there could not be electrons which did not obey the theory.

Such a conclusion may seem consonant with the thought, common to both empiricists and idealists (if worked out in different ways) that all necessity is projected onto the world by us. If the necessity of law is 'logical', it seems that it must be a function of our way of representing the world, rather than of the way the world is in itself.⁶⁶ Yet this conclusion is manifestly absurd. Nomic necessities and impossibilities constitute both the relations between means and ends and the hard limitations on action which we run up against in our existence in the world. It is absurd to suppose that we can be stopped from doing things and enabled to do others by the mere structure of our conceptual scheme or modes of representation. Moreover, if one looks again at the argument just presented for regarding it as logically necessary that gold is soluble in *aqua regia*, it will be seen that the argument itself presupposes that gold is so soluble of its nature, or essentially.

The dilemma can be resolved if we recognize that behind these 'logical' necessities lie less abstract constraints, and that we are here concerned with a deep relationship between our representation of the world and the world itself. One might say that there are natural necessities and impossibilities which generate 'logical' ones. In other words, we cannot sensibly think of nomic necessities as merely 'logical' or quasi-geometrical, *even though* propositions which state nomic relationships come out as 'logically' necessary. It is perhaps not surprising that reflection on the status of laws, coupled with the view that language as a whole is a 'theory', is one of the things which has encouraged the denial of any distinction between necessary and contingent propositions.⁶⁷

Rationalism in this form may in some respects resemble Locke's mechanism, but it is ultimately very different. That is because it does not start from Locke's conceptions of actuality and potentiality, essence and operation. Mechanism (or, at least, 'pure' mechanism) was founded on the principle that the ultimate properties of material things, from which their powers could in principle be deduced, are fully actual and, as such, fully perspicuous to the mind. For Descartes such perspicuity was perspicuity to the intellect, reaching only to extension and its modes. For Locke, it was perspicuity to sense and imagination: whatever the underlying properties of bodies may be, they are in principle perceptible as they are in themselves, if not by human senses, then by angelic ones. He took it that solidity possesses the credentials of a fundamental property just because he felt able to identify it as a simple sensible quality distinct from impenetrability, the power which depends on it. Now, whatever is wrong with such a view, the principle that potentiality is reducible to actuality, and that things cannot differ solely in power, is hard just to set aside. Yet in its interpretation there is room for a distinction among powers. It is intermittently operative powers, attributed in an *ad hoc* and untheoretical way, which call out for reduction. Such a property as Newtonian mass, on the other hand, which by Locke's (and, for that matter, Newton's) criteria is irremediably a power, is operative not intermittently, but unceasingly. In Newton's world there is no body which is not continuously affecting, and being affected by, other bodies in accordance with the law of gravity. It seems to be for just that reason that the mechanist's presumption that behind

every power lies something more actual, behind operation lies essence, seems doubtful in its application to the case of gravity and mass.

Historically, mechanism was attacked from two main directions, epitomized by Hume and Leibniz. Hume's arguments about causality and 'the modern philosophy' appear at their destructive best when they are read as a critique of just Locke's form of mechanism. He showed that, if we adopt Locke's understanding of actuality as perspicuity to the senses, no power or efficacy could be supposed to flow necessarily from the actual properties of things. Locke's attempt to derive impenetrability from solidity is summarily dismissed. The idea of solidity is the idea of a power, and it is absurd to suppose that the feeling of solidity resembles whatever it is that makes solid bodies solid.⁶⁸ As far as our conceptions go, things are in themselves inert, related only by contingent regularities. Given Lockean actual properties, anything may happen. Twentieth-century admirers of Hume have concluded that anything can happen, but that conclusion only seems to follow if we are not prepared to question Locke's notion of perspicuous actuality. Here we should remember the line of argument initiated by Leibniz and leading to a more dynamic conception of the actual and fundamental properties of material things. To adopt such a position is not to reject, but to reinterpret the metaphysical or ontological principle which mechanism presupposed.

Nevertheless, untenable as it may be, classical mechanism raises a further interesting question relating less to the principle at its heart, its penchant for actuality, than to the intuitive force of its paradigm of intelligible change. *Why* is it that mechanical interactions appear more intelligible to us than other experienced physical changes? A Humean is duty-bound to argue that it is simply a matter of familiarity, but that answer is utterly unconvincing. We are not much less used to the effects of heat and light than to the effects of the mechanical properties of things. Yet we do not immediately understand why heat will harden an egg as we can immediately understand why Arnauld's axle will turn one mill-stone but not the other. Greater familiarity with egg boiling, although it may reduce our wonder, will do nothing, without theory, to increase our comprehension. One relevant difference, perhaps, is that a certain practical, pre-theoretical mastery of mechanical causation, of how coherent bodies fit together and push one another around in space, is involved in our first identification of any independent objects at all. Yet that hardly explains why a sphere is so manifestly fitted for rolling.

The existence of God

The sketch towards a demonstration of the moral Law of Nature which Locke offered in the *Essay* assumed the possibility of demonstrating the existence of an omnipotent and wise Creator. Such a proof was proposed in *Essay* IV.x, a chapter which also constitutes an important text for the interpretation of Locke's view of laws of nature in the other sense, in particular of those laws linking thought with events in the brain.

I have argued that Locke's famous speculation of thought 'superadded' to a 'system of matter suitably disposed' does not, as Leibniz (if only polemically) assumed, envisage brute, arbitrary correlations which in some metaphysically mysterious, even quasi-miraculous way have managed to stick an incorporeal accident onto a material subject. Rather, he was proposing that it may be, for all we know, within the non-miraculous and intelligible potentialities of matter, when organized by a supreme intelligence, to constitute a thinking thing in something like the way in which it can constitute a living thing, a plant or an animal (the latter in any case possessing sensation). Passages from the argument for God's existence have often been taken in a sense directly contrary to this reading, in that they express rather emphatically the principle that 'bare incogitative matter' could never by itself constitute a thinking, intelligent being. Yet the intention of this principle, as I will endeavour to show, was simply to draw out the significance of a condition of our supposing a system of matter with the power of thought: namely, that we also suppose that that complex, organized body should have been organized by a supreme intelligence.

The power of thought is not, Locke held, a 'natural' property of matter in that it is not necessary to matter, i.e. it is not a 'property' of matter in the strict sense, flowing from its essence. Although it cannot be ruled out that some material mechanisms should think, we do at least know that the power of thought could only be an accident, rather than a property of body. It falls, as Locke put it, outside 'the bare essence or natural powers depending on the essence of matter in general'.⁶⁹ It was commonplace in logic to speak of accidents as 'superadded' to the subject. But the notion of superaddition by God tended also to have teleological or evaluative overtones both for Locke and his contemporaries.⁷⁰ An attribute 'superadded' by God or ascribed to his 'positive will' is not, in Locke's language, just any accident. Its acquisition is not a merely logical 'addition' but something expressly aimed for in God's plan, something not 'accidental' in the modern sense. Here it could, of course, be argued that everything which results from God's initial act of creation must have been not only foreknown, but intended by him. Not just the existence and initial state of the mechanical universe, but every element in its history from first to last, is equally the object of an initial complex intention. Locke, however, thought within the terms of a less holistic natural theology. The division of

labour between God and nature is, on his view, such that an initially chaotic state of particles in random motion, once created by God, might naturally have given rise to similar states of affairs without any further effort or direction on his part, so to speak. Yet there are certain ordered features of the world which could only have arisen from an initial state which (whatever its appearance) was carefully selected rather than chaotic, by motions which were initially directed rather than random. In other words, God must be supposed carefully to have chosen means to such complex and valuable ends. It is in general what Locke calls 'excellencies' or 'perfections' which are said to have been 'superadded' by God.

Locke tailored his notion of a 'perfection' with elaborate care in order that, unlike the notion employed by some other philosophers, it should be compatible both with his agnosticism about the natural world and with his mechanistic rationalism. The evidence for that careful and highly significant concern lies partly, but not entirely, in the argument for God's existence and in the peculiar relationship of that argument to earlier arguments, in particular to one occurring in Ralph Cudworth's *The True Intellectual System of the Universe*. Locke's debts here are well known; but the fundamental and surely deliberate difference from Cudworth has generally escaped all comment.⁷¹ It is that difference which helps to explain Locke's notion of a perfection, and its significance.

Cudworth's book is a detailed examination and criticism of ancient and modern atheism, which he ascribed to philosophical materialism. He advanced a number of arguments for God's existence, including the 'first mover' argument that matter cannot put itself in motion. Gravity, too, is taken as evidence of some agent in the universe over and above matter governed by mechanical principles. The chief argument, however, like Locke's, combines the traditional cosmological argument for an eternal first cause with an argument that the eternal cause of thinking things must itself be a thinking thing. The latter is presented as nothing but a special application of the former.⁷²

The relevant part of Cudworth's long argument is briefly as follows. First, since nothing can come from nothing, and something exists, then something must exist from eternity which is *causa sui*, existing necessarily or essentially (the standard 'cosmological' proof). So far, Cudworth thought, theists and atheistic materialists are undivided. But the atheists identify the eternal being with matter, arguing that the Lucretian principle that nothing can come from nothing rules out the creation of matter by something else. For the atheist, nothing can be created and everything is matter variously constituted. The correct application of *ex nihilo nihil fit* to creation, however, does not, Cudworth argued, rule creation out altogether, but merely restricts its possibility according to the directly derivable subordinate principle that 'no effect can possibly transcend the power of its cause'. The creation of one substance by another does not infringe this principle provided that the creator is higher on the chain of being than the creature: 'There is a scale, or ladder of entities and perfections in the universe, one above another, and the production of things cannot possibly be in way of ascent from lower to higher.'⁷³ Now experience assures us that there are finite thinking things. Thought has a higher place on the scale of perfections than any attribute of matter. The eternal creator of finite spirits must therefore possess thought (or, presumably, an even higher perfection). It must, moreover, possess such a perfection essentially—a point which independently

rules out the possibility that the eternal cause of animal and human consciousness should be matter:

Wherefore, if matter, as such, have no animal sense and conscious understanding, essentially belonging to it, (which no Atheists as yet have had the impudence to assert;) then can no motion or modification of matter, no contexture of atoms, possibly beget sense and understanding, soul and mind; because this would be to bring something out of nothing, in the impossible sense, or to suppose something to be made by itself without a cause. ⁷⁴

This argument has two connected features which should be emphasized. First, Cudworth's scale or ladder of perfections embodies a conception of sharp metaphysical differences in kind: it is a scale of different types of substances. Secondly, the argument relies on establishing such a difference between matter and spirit: if the argument is to work, the finite soul must be immaterial, and must be known to be so, if only from the evident superiority of thought on the scale of perfections. Cudworth was himself a corpuscularian and took real qualities and substantial forms to be 'long since justly exploded by the ancient Atomists, and expunged out of the catalogue of entities'. ⁷⁵ Consequently he assumed that whatever attribute can be thus reduced to mechanical modes of matter cannot claim for its possessor a higher place than matter in general on the scale of entities. But the reduction or 'resolution' of material forms and secondary qualities into 'mechanism and fancy' presupposes sensation and imagination, and it is self-evident that these cannot be themselves 'resolved' in the same way: 'wherefore undoubtedly they are no modes of matter or body, but attributes of another kind of substance incorporeal'. ⁷⁶ Colour in the object is reducible to mechanical attributes only if colour sensations can be kicked upstairs. Another argument for immaterialism is Cartesian: a mode cannot possibly be conceived without the substance of which it is a mode, but the 'life of the soul [and] cogitation may be clearly apprehended without body'. Once we have established the ontological distinction between mind and matter, there can be no doubt about their ordering on the scale of entities: there is 'a higher degree of entity, in minds and souls'. ⁷⁸ On the reasonable assumption that animals have sense but not intellect, there is a similar step up on the scale of entities from merely sensitive souls to souls with both sense and understanding.

It might seem surprising that Locke, who could be so carefully agnostic on the issue of materialism, should nevertheless have felt able to make use of Cudworth's argument. What he did, however, was almost tacitly to modify it in an absolutely characteristic way precisely in order to make it consonant with his agnosticism. The move he made is closely analogous to moves he made more explicitly elsewhere, e.g. in his account of miracles and in his treatment of personal identity. Miracles stand out by being contrary to normal experience, rather than to laws of which we have no knowledge. Personal identity consists in continuity of consciousness, rather than in the supposed identity of an unknown substance. ⁷⁹ Similarly, it is for Locke possible for us to identify and rank 'perfections' without knowing or caring what fundamental substances underlie them. Whether attributes are modifications of the same substance or of different substances will

simply be irrelevant to their ordering on the ladder of perfections. For Locke, even if human thought is a complex mechanical attribute of material machines, it is not thereby reduced on the ladder of perfections, as it would have been for Cudworth, to the level of all other attributes of matter: motion in general, solidity, extension, and so forth. Locke's scale is not really a metaphysical scale of 'entities' or 'being' (despite IV.x.4), but is simply a scale of valuable and remarkable attributes without regard to the general types of substances in which they are instantiated.

Such a notion of a perfection, and of degrees of perfection, is quite explicit in an argument in the second *Reply to the Bishop of Worcester* which has directly to do with the possibility of thinking matter, rather than with God's existence. As far as we are concerned, Locke argued, thought is an excellency like any other: a valuable and remarkable accident and therefore 'superadded', but not any kind of sure sign of a difference of substantial type. This is an argument so full and unequivocal that it is worth quoting from it at length:

For example, God creates an extended solid substance, without the superadding any thing else to it, and so we may consider it at rest: to some parts of it he superadds motion, but it still has the essence of matter: other parts of it he frames into plants, with all the excellencies of vegetation, life, and beauty, which are to be found in a rose or a peach tree, etc., above the essence of matter in general, but it is still but matter: to other parts he adds sense and spontaneous motion, and those other properties that are to be found in an elephant. Hitherto it is not doubted but the power of God may go, and that the properties of a rose, a peach, or an elephant, superadded to matter, change not the properties of matter.... But if one venture to go one step further, and say, God may give to matter thought, reason, and volition, as well as sense and spontaneous motion, there are men ready presently to limit the power of the omnipotent Creator, and tell us he cannot do it; because it destroys the essence, 'changes the essential properties of matter.' To make good which assertion, they have no more to say, but that thought and reason are not included in the essence of matter. I grant it; but whatever excellency, not contained in its essence, be superadded to matter, it does not destroy the essence of matter if it leaves it an extended solid substance; ...and if every thing of greater perfection, superadded to such a substance, destroys the essence of matter, what will become of the essence of matter in a plant, or an animal, whose properties far exceed those of a mere extended solid substance? ⁸⁰

This passage reveals that an important weapon in Locke's agnostic arsenal was the question of the status of animals. As in the discussion of identity, he structured his argument around the Aristotelian trichotomy of vegetable, animal and human life or souls. He assumed that, although animals have sensation and self-motion, a materialist account of the first two forms of life can be given. Yet Cudworth, who favoured animal souls (possibly recycled), regarded it as established that either animals are bare machines without sensation, as the Cartesians said, or they have immaterial substantial souls. The

strength of Locke's position was that both views might well seem eccentric and arbitrary, while the latter (if immateriality and immortality are supposed to stand or fall together) might be thought to have theologically unsound consequences.⁸¹ If, as Locke supposed, Stillingfleet accepts neither, he must admit that 'God can and doth give to some parcels of matter a power of perception and thinking'.⁸² But it is also clear that, as Locke saw it, even vegetable life, although nothing but the continuous operation of a suitable 'organization of parts', is as much a superadded or superinduced 'perfection' or 'excellency' as is animal life or human consciousness. He took none of these mechanical perfections of certain parcels of matter (supposing that that is what they all are) to be explicable by the essence of matter alone. As in the case of motion, since these perfections are not possessed by all material (i.e. solid and extended) things, it follows that they are not possessed necessarily by any. They are not 'properties' of matter deducible from its underlying essence alone, whatever that may be. Equally, none of them is known to be contrary to the essence of matter.

This deviation from Cudworth's notion of a perfection covertly changes the whole character of the argument for God's existence. The similarities with Cudworth are, by contrast, obvious. The 'cosmological' first step remains unaltered. Locke next appeals to the seemingly Cudworthian principle that

whatsoever is first of all Things, must necessarily contain in it, and actually have, at least, all the Perfections that can ever after exist; nor can it ever give to another any perfection that it hath not, either actually in it self, or at least in a higher degree.⁸³

Gestures are made towards deriving this principle from *ex nihilo nil fit*.⁸⁴ Given the premise that a thinking thing exists (whether material or immaterial is beside the point), it follows that the eternal first cause is a 'cogitative Being'. Then follows an argument that the eternal cogitative being which is the cause of cogitation in us could not be material, since if it were its cogitation would be unexplained, and arise from nothing. In other words, as in Cudworth, the eternal cause of thought must itself think essentially or necessarily. This last argument must be examined carefully, since, if the present line of interpretation is correct, it was more important to Locke than to Cudworth, for whom 'thinking matter' was in any case a palpable contradiction. Moreover, parts of Locke's exposition seem on their own open to a natural and generally accepted interpretation which is quite contrary to what Locke himself says elsewhere, and even to his explicit account of this very argument's meaning. It is here, moreover, that the divergence from Cudworth is revealed. Let us, then, focus on this argument.

Both Cudworth and Locke were arguing that, if the possession of consciousness by the eternal first cause is itself to have a cause (as it must), then it must derive from the essence of that eternal being. Even Cudworth at this point was not so much arguing directly that no material thing could think as that the eternal first cause could not of itself bring about thought unless, unlike matter, it thought essentially. Its thought, like its existence, must have a source, and *ex hypothesi* there is no outside source. Both philosophers linked thought with motion. First it was stated that motion does not flow

from the essence of matter. The motion in existence must therefore have had an external efficient cause.⁸⁵ Despite this conclusion, the impossible supposition was then accepted for the sake of argument that matter is eternally (and necessarily) in motion. On that supposition, could its motion have been the source or cause of thought? Here Cudworth made it plain that his negative answer rests on the presumption that for matter to 'beget sense and understanding' would be for it to beget 'soul and mind', a new and different substance. For example, he agreed that, if thought could ever be even an accident of matter, matter's coming to possess it would not contravene *ex nihilo nil fit*, since there would be 'no real entity produced out of matter, which was not before in it, but only new modifications'.⁸⁶ For Locke, however, the principle that consciousness cannot arise from mere matter-in-motion seems to have had a rather different significance.

Locke argued, not that to postulate thought is to postulate a substance superior to matter, but that no perfection can arise even as a mechanical modification of matter if it neither flows from the essential nature of matter (or, for the sake of argument, of matter-in-motion) nor has an appropriately powerful external cause. If matter supposed necessarily in motion would not itself *necessarily* produce a certain perfection, then it could not do so by itself *accidentally*. Indeed, an excellency or perfection of matter is superior both to the 'properties' of matter and to random or undirected accidents (i.e. to what happens to matter 'accidentally' in something like the modern sense) precisely in that it could not possibly have arisen by bare chance. The generation of thought thus implies an antecedent thinking thing just because it implies an intelligent purpose and skill. That explains why the acme of the scale of entities, in possessing perfect understanding, possesses other perfections 'eminently': it is equipped to produce all lesser perfections without actually possessing them.⁸⁷ The creator of rational beings must be rational, but the creator of vegetables, provided that it is rational, need not vegetate. Such a connection between teleology and the principle that an effect cannot transcend its cause was not novel. It had appeared, for example, in Descartes' illustration of his notorious epistemological principle that the cause of an idea must be at least as far up the scale of being as its intentional object: his instance was the thesis that the cause of an idea of a machine must be, if not its object actually existing, then the inventive intelligence of its possessor, who is ontologically superior to the machine.⁸⁸ Indeed, a connection with teleology had been present in the Aristotelian version of the principle, although here the idea is cause rather than effect: man begets man, but the idea of a plough in the mind of a craftsman is the cause of a plough. Theology made the model of the craftsman comprehensive. In effect, Locke had converted Cudworth's extended cosmological proof, which is about substantial creation, into a deductive form of the argument from design; or at least into a conflation of the cosmological and ideological proofs. He supposed it not merely deductive, but also peculiarly a 'demonstration' just because, whereas the standard argument from design can start from any purported 'evidence' in observable nature, this argument starts from the one such premise known with intuitive certainty from one's own case, the premise that a thing with 'sense, perception and knowledge' exists. Moreover, it appeals to the highest perfection in our experience, one directly analogous to the intelligence requisite in the divine architect.

There are several passages in IV.x which indicate that Locke was well aware of this

transformation. His rhetorical gloss on the whole argument, aimed at those unconverted by its first bare formulation, is as follows: it is ‘senselessly arrogant...to suppose Man alone knowing and wise, but yet the product of mere ignorance and chance; and that all the rest of the Universe acted only by that blind hap-hazard’. He quoted Cicero: ‘What can be more sillily arrogant...than for a Man to think...that those things, which with the utmost stretch of his Reason he can scarce comprehend, should be moved and managed without any Reason at all?’ Again, he criticized those who, like Descartes, preferred to argue from the idea of a perfect being, for trying to undermine ‘those proofs, as being weak, or fallacious, which our own Existence, and the sensible parts of the Universe, offer so clearly and cogently to our Thoughts, that I deem it impossible for a considering Man to withstand them’.⁸⁹

It is Locke’s placing his argument within the general context of the argument from design which should colour our interpretation of his often and variously repeated claim that ‘it is as impossible to conceive, that ever bare incogitative Matter should produce a thinking intelligent Being, as that nothing should of it self produce Matter’.⁹⁰ It is significant that he more than once suggested that what is impossible is the *genesis* of thought in a postulated eternal matter, even though the point is not strictly concerned with a temporal beginning but rather more generally concerns the causal source or origin of thought, which ought indeed to lie timelessly in the essence of the eternal thinking thing: ‘I appeal to everyone’s own Thoughts, whether he cannot as easily conceive Matter produced by *nothing*, as Thought to be produced by pure Matter, when before there was no such thing as Thought or an intelligent Being existing.’⁹¹ Here is another passage:

if we will suppose nothing first, or eternal; *Matter* can never begin to be: If we suppose bare Matter, without Motion, eternal; *Motion* can never begin to be: If we suppose only Matter and Motion first, or eternal; *Thought* can never begin to be. For it is impossible to conceive that Matter either with or without Motion could have originally in and from it self Sense, Perception, and Knowledge, as is evident from hence, that then Sense, Perception, and Knowledge must be a property eternally inseparable from Matter and every Particle of it.⁹²

Now this argument, as it stands, is clearly neutral on the question whether *finite* ‘cogitative beings’ are in the nature of things immaterial. Just as the being and motion of matter, because they do not derive from the essence of matter, require an external cause or source, so thinking ascribed to matter supposed for the sake of argument eternal and eternally moving would require an external cause and source. Like them, it cannot be conceived of as flowing from the essence of matter—even if, for the sake of argument, we assume that they do. To deny that matter can have sense and reason ‘originally in and from itself, i.e. essentially or necessarily or, in Locke’s special sense, ‘naturally’, is not to deny that a material thing can have them *tout court*, or ‘naturally’ in a more general or popular sense of the word. The external cause must be an architect, but need not perform metaphysical miracles.

It is likely to be objected to these suggestions that Locke clearly stated in the course of the argument that it is impossible that the motion of matter should *ever* constitute

thought. Indeed, it might be claimed that Locke simply concluded that the eternal cogitative being cannot be material from the premise that no cogitative being can be material. There are two passages in particular which can seem to bear out this contention. Immediately before the passage last quoted, Locke had written:

you may as rationally expect to produce Sense, Thought, and Knowledge, by putting together in a certain Figure and Motion, gross Particles of Matter, as by those that are the very minutest, that do any where exist. They knock, impell, and resist one another, just as the greater do, and that is all they can do. ⁹³

Yet this is part of an argument that thought cannot ‘be produced by pure Matter, when before there was no such thing as Thought or an intelligent Being existing’. It is immediately followed by the analogy with being and motion, the second edition gloss which emphasizes that the impossibility concerns matter’s possession of thought ‘originally in and from itself, and the separate argument for the same conclusion based on the undoubted existence of some unthinking matter. It is reasonable to assume that Locke was here arguing for what he said he was and for nothing else. In that case his argument was that, however we pore over the (presumed) essential properties of matter, we shall never see how sensation and thought will ‘flow’ from them alone: left to itself, the undirected or random motion of particles merely lumped together will produce nothing but more of the same. The point of the sentences quoted immediately above is very specific: in effect, that it will not avail the atheist to argue that motion at the level of the finest minute particles may for all we know necessarily produce thought, since mere size introduces no differences in principle. What is lacking from his story is suitably potent external direction.

The second passage which raises the same issues runs as follows:

For to suppose the eternal thinking Being, to be nothing else but a composition of Particles of Matter, each whereof is incogitative, is to ascribe all the Wisdom and Knowledge of that eternal Being, only to the juxta-position of parts; than which, nothing can be more absurd. For unthinking Particles of Matter, however put together, can have nothing thereby added to them, but a new relation of Position, which ‘tis impossible should give thought and knowledge to them. ⁹⁴

These sentences, as they stand out of context, are also, of course, open to interpretation as the claim that even God cannot make a material system (‘however put together’) think, at any rate non-miraculously. Yet it can equally well be interpreted as what in its context it purported to be, an argument that, if *per impossibile* the eternal, wise, and knowledgeable God were a system of matter, then his wisdom and knowledge, not to speak of his being, would be unaccounted for: they could be ascribed neither to a suitably endowed outside source nor to his essence, but only to a chance relationship between unthinking parts, which is absurd. The phrase ‘however put together’ must be read in the context of the assumption of the argument that the parts have been put together accidentally or

randomly. It is the possibility that the chance or undirected relationship between parts should constitute the thought of the eternal being which is ruled out by the principle that the effect cannot transcend the cause. Locke's argument continued with the point that, if the material God's thinking 'consists in a certain motion of the parts', then his thoughts

must be unavoidably accidental, and limited; since all the Particles that by Motion cause Thought, being each of them in it self without any Thought, cannot regulate its own Motions, much less be regulated by the Thought of the whole...So that such a thinking Being will be no better nor wiser, than pure blind Matter; since to resolve all into the accidental unguided motions of blind Matter, or into Thought depending on unguided motions of blind Matter, is the same thing.⁹⁵

This further argument at least leaves it open that an immaterial and far from blind supreme mechanic should so regulate and guide the motions of the minute parts of a material mechanism that, out of his own regulative wisdom and knowledge, he bestows wisdom and knowledge on the system. If we suppose such a divine external cause, the problem of how *unguided* and *haphazard* motions of the incogitative parts of a system of matter should constitute its thought is circumvented, since the relevant motions are guided or selected: while the ideological principle that the effect must derive its being or perfection from the cause is neatly illustrated. All this explains what Locke meant in talking of the (as far as we know) possible 'superaddition' of thought to matter. Yet in all this no contranatural or supermechanical operations of matter are postulated. For Locke, God must have guided or ordered his material creation, but his performance of the metaphysical absurdity which commentators often understand today by the term 'superaddition' was never envisaged. Such an absurdity played no part in the materialism with which Locke was familiar and to which he was directing his response. To sum up that response: Hobbes might be right about us, but his notorious implication could not be right that God, too, is material.⁹⁶

If this interpretation of some sentences on the page seems to rely too much on the employment of context to subvert their natural interpretation, there is nevertheless independent evidence that it embodies Locke's own view of his intentions. In the face of Stillingfleet's criticism, including the imputation that the *Essay* contradictorily asserts both that we can know that what thinks in us is an immaterial spirit, and that we cannot know that it is not material, Locke's defence contained this categorical statement:

...from the idea of thinking, we can have a certainty that there is a thinking substance in us; from hence we have a certainty that there is an eternal thinking substance. This thinking substance, which has been from eternity, I have proved to be immaterial. This eternal, immaterial thinking substance, has put into us a thinking substance, which, whether it be a material or immaterial substance, cannot be infallibly demonstrated from our ideas; though from them it may be proved, that it is to the highest degree probable that it is immaterial.⁹⁷

On the present interpretation of *Essay* IV.x, this is just what we should expect Locke to have said, given that he is in general prepared to grant that the ‘inconceivability’ or incomprehensibility to us of a material mind creates at least some prejudice in favour of dualism⁹⁸ (despite the bolder allegation at IV.iii.6 that dualism is no more conceivable than materialism). It might be argued, however, that another explanation of Locke’s words is possible: he might have meant that the intrinsic or natural impossibility of a material system’s thinking rules out God’s being material, but that an immaterial God has it in his power to do what is intrinsically or naturally impossible (e.g. to make a material system think), although it is more probable that he would stick to what is possible. It seems a strange and tortuous line of thought, however, to postulate that God is bound to the possible in his *being*, but not in his *doing*. It could not reasonably be ascribed to Locke unless he had indubitably maintained elsewhere that the superaddition of thought to material systems would be contranatural and quasi-miraculous. Yet, on the contrary, there is extremely strong evidence that he did not think this. We are told that thought would need just as much to be ‘superadded’ to immaterial substance as to material substance; that thought is, as far as we know, a perfection no different in kind from animal or vegetable life; that dualism is no more conceivable than materialism; that to assume that what is inconceivable is therefore impossible is to limit God’s power to the extent of our ideas; that no one has proved a contradiction in the notion of thinking matter; and so forth. Ironically, one reason why these points have been neglected or misinterpreted is probably that commentators have felt secure in their belief that *Essay* IV.x.10 and IV.x.16 unequivocally assert the impossibility of thinking matter.

The changes made to the later editions of the *Essay*, presumably for the purpose of clarification, do more to bear out the explanation here proposed. In the second edition, for example, IV.x.10 came heavily to emphasize that the issue is simply whether matter could have sense and reason ‘originally in and from itself’, i.e. essentially or necessarily or, in Locke’s technical sense, ‘naturally’. Similarly, IV.iii.6 came to state that, although it is shown in IV.x to be a contradiction that the ‘Eternal first thinking Being’ is material, it is not a contradiction in general that material systems should think.⁹⁹ Another passage, II.xxiii.15, is an argument expressed within the terms of dualism (to the effect that our idea of spirit is as clear as our idea of matter, although we have no positive idea of the substance of either) which can misleadingly seem like a would-be demonstrative argument for dualism. The fourth edition version, constructed at a time when Stillingfleet’s charge of inconsistency would have been fresh in Locke’s mind, ends with the sentence, ‘[Sensation] I must be convinced cannot be the action of bare insensible matter; nor ever could be without an immaterial thinking Being.’¹⁰⁰ This sentence only becomes intelligible if taken to be an allusion to IV.x, contrasting ‘bare incogitative’ or ‘blind’ matter with matter guided by a ‘thinking eternal Being’. Given all this, it is not easy to attribute to Locke the view that the contradiction in the notion of a material God simply derives from a contradiction in the notion of a material thinking thing. Whatever others might have meant, or whatever we might now mean, in uttering sentences like certain ones selected from IV.x, it seems as sure as interpretation can be that this was not Locke’s meaning.

It cannot be denied, however, that the argument for God’s existence could have taken a

form more obviously consonant with the strenuous agnosticism of IV.iii.6 and elsewhere; nor that the total response to Stillingfleet contains an unacknowledged element of fudging or tidying up. For example, IV.x is one of those passages in which Locke seems ready to assume that Boyle's account of the essence of matter will do, and to forget that 'solid, extended substance' ought by his own lights to be regarded as no more than a definition of our *idea* of matter, equivalent to a nominal essence. Once Boyle's mechanism is subjected to the sceptical argument of, for example, II.xxiii.23f (not to speak of the problem of gravity), the claim that, without 'an intelligent being existing', particles in motion 'knock, impell, and resist one another... and that is all they can do', even if it has the meaning here attributed to it, can only seem to rely on an unduly dogmatic assumption as to the essence of matter. It is therefore out of place in a purported *demonstration*, in which every lemma must be absolutely certain. Moreover, the independent argument from the premise that some matter evidently does not think to the conclusion that matter does not think essentially, however reasonable, is an appeal to sense experience and so less than is strictly required for a demonstration.¹⁰¹ What Locke therefore needed was a general *a priori* principle which is independent of a specific mechanist theory, i.e. the principle that, *however* particles in motion operate in general and as such, these operations, *unguided or unselected by thought*, will not lead to thought. Locke would, of course, have accepted such a principle, deriving it from the extended cosmological principle that the cause must have at least the perfection of the effect. Thus rewritten, Locke's argument would be less graphic, but more evidently unreliant on a specific, dogmatically assumed form of mechanism.

It is interesting that Locke himself seems to have recognized the need for just such a revision of the argument of the long passage quoted above from the *Second Reply*.¹⁰² There he argued that the superaddition of thought to matter no more evidently 'destroys the essence of matter', i.e. extension and solidity, than does the superaddition of any other perfection. Yet that was less than clearly to recognize that, by his own lights, 'extended, solid substance' is only a nominal, rather than the real definition of matter.¹⁰³ He seems rather to have been assuming the corpuscularian view of the latter for the sake of argument. Yet in a later letter to Collins the same general point was put in a way which is more scrupulously compatible with agnosticism about the ontological essence. Cogitation, extension and solidity, it is there said, may all be, for all we know, similarly related to an unknown substance or essence: 'Of this substance we have no idea, that excludes cogitation, any more than solidity.'¹⁰⁴ If only Locke had seen fit to rewrite the relevant parts of the argument for God's existence in the same terms, then perhaps fewer mistakes would have been made about his meaning.

Philosophical criticism of Locke's proof can be brief, but it may help us to grasp its character if we identify its weaknesses. Crudely, it is either invalid or circular.

An argument from design is normally neither presented nor to be judged as a would-be deductive proof. As Cudworth, for example, put the teleological argument in its most general form, it is no more possible that the regular system of the world should be the result of the fortuitous motion of atoms than that the hated 'six books of T.Lucretius Carus' should be a chance arrangement of letters.¹⁰⁵ Such things were held to be 'impossible' because of the supposed massive improbability of their happening by

chance, rather than in the sense in which a triangle with angles greater than two right angles is impossible. Locke, however, tried to conflate the two sorts of 'impossibility' by means of the extended cosmological principle. Thus, from premises which we can know to be true by observation and experience (e.g. that living things exist, or, more certainly and more impressively, that a thinking thing exists) we can, so Locke believed, strictly deduce the existence of God. Yet (whatever we are to make of his cosmological principle) the argument is invalid unless we presuppose the additional premise that the observable or introspectible attribute in question, life, thought or whatever, is a 'perfection' in the required sense: a point which is beyond the powers of observation or introspection to determine by themselves. For, in effect, as we have seen, the judgement that the attribute is a perfection just is the judgement that it involves too much remarkable organization to have come into existence by chance. The intuition to which Locke appealed in each case is really no different from the sort of intuition appealed to by Cudworth in the example of Lucretius' poem. To make explicit and to unpack the necessary additional premise is therefore to make explicit the circularity of the whole argument, and to reveal its inherent tendency to collapse into the standard probabilistic argument from apparent design, a tendency manifested in Locke's own rhetoric.

Locke himself brought an analogous charge of circularity against an argument employing the rival concept of a perfection. For in the same letter to Collins, he considered an argument of Norris's for substance-dualism, that 'Cogitation...is more excellent than motion, or vegetation; and therefore must belong to another substance than that of matter, in the idea whereof, motion and vegetation are contained.' The point of one criticism which Locke went on to make seems to be that Norris offered no independent criterion by which to identify a difference in excellence. Consequently he argued in a circle, judging thought to be superior to motion and vegetable life just because he already accepted the conclusion that matter cannot think, but can move and vegetate. Yet Locke himself could be embarrassed by the reasonable demand for an independent criterion for distinguishing 'perfections' from those manifold accidents of matter which count for no more than its undirected arrangement and motion: a criterion independent, that is to say, of the conclusion to be drawn that the former could only have arisen by design.

The Law of Nature and human freedom

The reader is introduced to the topic of ethics in the first two books of the *Essay* very much in the context of the doctrine that all our ideas are derived from experience. The derivation of value proposed is not roundabout. Its fundamental sources are pleasure and pain, simple ideas which accompany sense-perception and thought. The argument in effect assumes that all pleasure (or all pain) is the same in kind whatever its object, differing intrinsically only in degree. Bodily pleasures differ from pleasures of the mind only in their objects or causes, and in general that is how the various ‘passions’ (which Locke called ‘modes of pleasure and pain’) are distinguished from one another. Things are *good* for the individual in so far as they promote his pleasure and diminish his pain (i.e. promote his happiness), and *evil* as they do the reverse.¹⁰⁶

The pursuit of happiness is limited or, rather, directed by our conceptions of right and wrong. Where later utilitarians have been inclined to define the ‘right’ course of action simply as the one with the best foreseeable consequences, Locke proposed a rather less direct connection between ‘good’ and ‘ought’. What is ‘right’ or ‘wrong’ is what is in accordance with, or contrary to, law. The idea of law involves the idea of a law-maker with the right to legislate and the power to enforce the law with punishments and rewards. ‘*Morally Good and Evil*, then, is only the Conformity or Disagreement of our voluntary Actions to some Law, whereby Good or Evil is drawn on us, from the Will and Power of the Law-maker.’¹⁰⁷

There are, according to Locke, three kinds of law. First, there is ‘that Law which God has set to the actions of Men whether promulgated to them by the light of Nature, or the voice of Revelation’. Such a law is ‘the only true touch-stone of *moral Rectitude*’. *Knowledge* of its content rests with ‘the light of Nature’, since, as it was seen in Volume I, even direct revelation cannot for Locke give rise to knowledge. Knowledge is possible, it seems, because we can know that the law is benevolent: ‘He has Goodness and Wisdom to direct our Actions to that which is best.’ Second, there is the civil law, which determines what is criminal. Third, there is ‘the law of *Opinion* or *Reputation*’, which ‘by a secret and tacit consent’ determines the ordinary notions of the virtues and vices, the subject-matter of non-theistic moral philosophy. This last law (called in the first edition the ‘philosophical’ law)¹⁰⁸ may differ in some respects from society to society, but there exists wide agreement with respect to its central core, as well as a broad correlation between that core and the Law of Nature. That is because, when men join together in judgement, then out of self-interest they normally commend whatever evidently promotes the common good and tends to the preservation of society. Conventional morality thus has a Hobbesian basis which brings it fairly close to true morality in content, if not in spirit. It is the praise and blame themselves which constitute the sanctions of the law of

reputation.

Of human laws we can ask by what right they command obedience, and an answer to this question must, in Locke's view, bring us back to natural law, or moral law proper. A main theoretical purpose of the *Second Treatise on Government* was to set the conditions of the obligation to obey the civil law by reference to the divine law. Why then should *natural law* be obeyed? Locke held it self-evident that rational creatures have an obligation to obey their creator and sustainer:

The *Idea* of a supreme Being, infinite in Power, Goodness, and Wisdom, whose Workmanship we are, and on whom we depend; and the *Idea* of our selves, as understanding, rational Beings, being such as are clear in us, would, I suppose, if duly considered, and pursued, afford such Foundations of our Duty and Rules of Action, as might place *Morality amongst the Sciences capable of Demonstration*.¹⁰⁹

All this raises certain questions of interpretation. Some, which will not be pursued far here, concern the role of the so-called 'law of reputation'. This law primarily functions in Locke's argument as an explanation of the force of non-theistic ethics, whether everyday and intuitive or systematic and philosophical. His emphasis on the ultimate relativity of any such morality and its variability with variations in taste constituted a critique of attempts to arrive at a natural ethics without reference to God. His evident fascination with travellers' tales of other societies seems to have centred on the question of differences and similarities in moral concepts and moral codes. In effect his explanation of the level of agreement was at once both a concession to Hobbes (and others)¹¹⁰ and a criticism. Genuine morality has an authority antecedent to organized society, and cannot be the creation of man for the sake of social harmony. Yet Locke's attitude to the law of reputation was ambivalent. On the one hand it can conflict with the divine law,¹¹¹ but on the other he seems to have seen its general tendency to overlap that law in its content as more than the likely coincidence of a tacit covenant among self-interested agents with the dictates of a just and benevolent lawgiver. In 'Of Ethick in General', which is apparently the draft of an aborted chapter of the *Essay*, Locke suggested that the widespread conception of ethics as a science proper to philosophers stems from a sort of dim appreciation of the Law of Nature. It is only because philosophers have commonly failed to recognize moral rules as divine commands that they have been liable to appeal only to the sanctions of reputation and disgrace, while the content of secular ethical systems has been restricted either to advocacy of what tends to the preservation of society, or to mere definitions of the actions of which a particular society approves or disapproves (either Hobbes or Aristotle, so to speak).¹¹² A rather different connection between the two laws was suggested in *Some Thoughts on Education*, in which it was proposed that the child should be brought to a knowledge of ethics by 'the love of reputation, instead of satisfying his appetite, being made habitual to him'.¹¹³ One could indeed hardly fail to recognize the role in moral education of simple esteem and shame.

These thoughts about the relation between a divinely sanctioned Law of Nature and human or secular morality were in effect contributions to a standing theme of moral

philosophy in the seventeenth century. Secular or 'pagan' morality existed, whether as philosophically argued systems or commonly accepted principles of justice and social conduct. How then was it to be explained, and how were these supposed duties to our fellow men related to our duty to God as set out in the Bible or apprehended by natural reason? It seems beyond doubt that an important influence on Locke's thinking about this question was Nicole's *Essais de Morale*, three of which he admired sufficiently to translate, originally with a view to publication. In one of the latter, and elsewhere in the *Essais*, Nicole had elaborated a conception of the conformity between the duty required of us by faith and considerations of self-interested reason, or the 'rule of worldly wisdom', which enjoins us to treat others well in our own immediate interests and in the interests of society (and so again, ultimately, in our own interest). At the same time Nicole emphasized and deprecated the almost irresistible influence on us of the desire for the approbation of others, a passion which commonly leads us to abide by the divinely sanctioned principles of justice and 'civility' necessary to society, but which he seems to have regarded as inevitably self-defeating in so far as it is necessarily bound up with the less attractive manifestations of vanity and hurt pride. Locke did not take over the whole of Nicole's intricate analysis of the relations between self-interest, human approbation and God's will, but it was almost certainly the stimulus for his own conception of two intimately entwined moralities, human and divine.¹¹⁴

Another, very different question relates to the extent of the analogy between ethics and mathematics which dominates so many of Locke's allusions to morality in the *Essay*. That analogy, which evidently reflected his respect for the sort of reasoning about Natural Law to be found in such writers as Hooker, Suarez, Grotius and Pufendorf, had been present in his first extended exercise in moral theory, the manuscript now known as *Essays on the Law of Nature* and taken to have been composed in the early 1660s. Yet his conceptions both of ethics and of mathematics seem later to have undergone some significant refinement.

In both works Natural Law is 'knowable by the light of nature', i.e. reason operating on the materials supplied by the senses. In the *Essay*, however, these materials are simple ideas, whereas in the *Essays* they include not only 'ideas of particular sensible things' conceived of as given 'images' from which reason can construct further images, but also prepositional premises: for 'every argument proceeds from what is known and taken for granted, and the mind cannot discourse or reason without some truth that is given and received'. This latter model for the division of labour between reason and the senses was extended in the *Essays* to mathematics, ascribing knowledge of axioms to the senses: 'Surely mathematical reason (*mathesis*) presupposes these objects of its operations together with other common principles and axioms; it does not discover them nor prove them.'¹¹⁵ With respect to ethics two kinds of empirical premise are specifically identified: first, the premises of a teleological proof of the existence of a God to whom we owe creaturely duties; and second, premises about the created faculties and constitution which we find in ourselves, and from which specific duties are allegedly deducible. Duties to do what is necessary for self-preservation and the preservation of society, for example, can be deduced from the existence in us of natural inclinations both to preserve ourselves and to live in society. Such reasoning is explicitly distinguished

from the attempt to found the Law of Nature on a principle of self-interest in the manner of Hobbes. The latter confuses 'extraneous law' (*aliena lex*) with 'private utility' (*propria utilitas*), whereas the former considers human nature simply as an indication of the will of a wise and beneficent creator: for 'a great number of virtues, and the best of them, consist only in this: that we do good to others at our own loss'.¹¹⁶

By the time of the *Essay*, however, Locke had moved to a view of mathematics according to which, although mathematical ideas ultimately derive from experience, axioms are intuitively perceived relations between abstract ideas which are independent of sensitive knowledge. Correspondingly the role of empirical premises in ethics became at any rate less prominent. Indeed the very notion of 'demonstrative' moral knowledge entailed that any premises are known intuitively, rather than through the senses.¹¹⁷ Yet an existential premise, even if concerned with the existence of oneself and so, according to Locke, knowable intuitively, is still empirical or factual, going beyond our ideas. Moral knowledge must, it seems, rest on knowledge of the existence of at least one interpersonal relationship, between an omnipotent, wise and beneficent creator and his rational, free creature, capable of pleasure and pain. That alone would be enough to mark off moral science from mathematical, which is not concerned with existence at all. As we have seen, Locke often rested the analogy between them on the claim that moral truths about, for example, adultery or gratitude are independent of the *existence* of adultery or gratitude.¹¹⁸ The present point is that, on his account, it seems that there are no moral truths independent of the existence of a creator with a certain nature, and a creature with a certain nature.

Locke was evidently aware of the problem. Even in the *Essays* he had made a suggestion seemingly designed to remove existential propositions from ethics by making moral propositions conditional: he argued that it follows 'from the nature of man that, if he is a man, he is bound...to observe the law of nature, as it follows from the nature of a triangle that, if it is a triangle, its three angles are equal to two right angles'.¹¹⁹ Yet the distinction drawn in the *Essay* between a science of modes and a science of substances made it possible to object to the analogy between mathematics and ethics precisely on the ground that ethics rests on premises about substances, in particular about man himself. Locke responded to this foreseen criticism with the sophisticated argument that, although moral reasoning involves ideas of substances as well as ideas of mixed modes and relations, moralists construct and employ their idea of man in a way different from naturalists: 'when we say that *Man is subject to Law*. We mean nothing by *Man*, but a corporeal rational Creature: What the real Essence or other Qualities of that Creature are in this Case, is no way considered'. The naturalist's idea is inherently open to replacement in the light of fresh discoveries, but the moralist uses an 'immoveable unchangeable *Idea*'. The natural species of a 'moral Man' is irrelevant to his or her moral obligations, just as what a cube or globe is made of is irrelevant to geometrical reasonings about it.¹²⁰

Here Locke was moving further and more self-consciously away from the broadly Aristotelian conception of ethics as a systematic, quasi-biological study of man's nature as a rational, social species of animal, and towards a conception of ethics as the derivation of an *a priori* law for all rational, sensitive beings. No doubt this move helped

to make possible a theory like Kant's, but, if that thought is being pursued, it has to be stressed that for Locke, as not for Kant, a capacity for pleasure and pain is as essential a condition of moral agency and motivation as rationality. It is hardly less obvious that Kant was expressly denying Locke's claim that the derivation of the moral law requires reference to the divine lawgiver.

The analogy between the 'moral man' and mathematical objects came, as we shall see, to play a crucial and questionable role in Locke's thinking about 'persons'. The analogy is less than perfect in so far as the mathematical object can be said to be the abstract shape rather than the material or substantial cube or globe. A rational agent, on the other hand, is as such substantial. Or so it will be argued below.

Locke's emphasis on the philosophical need to bring God into moral reasoning makes it not entirely easy to answer what is perhaps the most obvious question raised by his ethics. Is his Law of Nature, like that of St Thomas or Suarez, intrinsically rational and necessary, an immutable and dependable thought of a wise and good God, and such that God himself is in some sense bound by it? Or is it, like Ockham's, a contingent and arbitrary decree, such that God might, at any rate in principle, change his mind or make exceptions without reproach? It might seem that Locke's position is close to the former, since, unless God's goodness were intrinsically dependable, the law could not be known. In the *Essays* it was asserted categorically that the Law of Nature does not depend on an unstable and mutable will, but on the eternal order of things, and it was in this context that moral principles were presented as conditional:

certain duties arise out of necessity and cannot be other than they are. And this is not because...God...could not have created man differently. Rather, the cause is that, since man has been made such as he is, equipped with reason and his other faculties and destined for this mode of life, there necessarily result from his inborn constitution some definite duties for him, which cannot be other than they are. In fact it seems to follow just as necessarily from the nature of man that, if he is a man, he is bound...to observe the law of nature, as it follows from the nature of a triangle that, if it is a triangle, its three angles are equal to two right angles, although perhaps very many men are so lazy and so thoughtless that for want of attention they are ignorant of both these truths.

Not only was the analogy with geometry already in place in this early passage, but we can see a link with Locke's later treatment of the laws of physics. The necessity of the law is hypothetical, but hard: God was free to will what laws he liked in that he was free to create what things he liked, but in creating free and rational beings capable of pleasure and pain he *ipso facto* willed a certain law for those beings; just as, in choosing to create matter, he chose certain necessary laws of motion. At the same time there is a significant difference between Locke's accounts of the two sorts of law, and of our knowledge of them. The ultimate, intelligible laws of motion, whatever they are, he took to flow from the essence of matter alone, whatever that is. In this, as we have seen, he differed from Descartes, for whom knowledge of the laws of mechanics requires reflection not only on the essence of matter, but on the immutable nature of God who wills them. There is no

proposal in the *Essay* that a science of the motion of bodies would need so to be based on theistic metaphysics. In the case of the moral Law of Nature, however, despite the passing claim in the *Essays* that moral duties flow from the nature of man, the crucial factor is not purely internal to man but relational: man is the creation of an omnipotent and beneficent God. So without reflection on at least some aspects of God and of our relationship to God, we cannot know our duty.

It is just that (as it might seem to us) particularly archaic feature of Locke's system which separates it not only from the Thomistic tradition, but from what might be regarded as the theoretically more progressive rationalistic ethics of the seventeenth century, advanced in their different ways by such philosophers as Grotius, Cudworth and Samuel Clarke.¹²¹ The point of divergence might seem not only archaic but gratuitous, since there is no sign that Locke was attached to the theological doctrines characteristic of strong, seriously meant voluntarism. He found the doctrine of original sin, deserving 'endless torment, in hell-fire', 'little consistent with the justice or goodness of the great and infinite God'. He suffered, it seems, no conceptual embarrassment in applying moral epithets to God himself in their plain sense, or in supposing God bound, without any peculiar promise or arbitrary covenant, by the dictates of morality and justice as we conceive of them.¹²² Moreover, as it has been argued at length above, he by no means preferred revelation to natural reason as the source of moral knowledge. If in *The Reasonableness of Christianity* he came to assign revelation greater importance for the moral life than he had done in the *Essay*, that seems to have been only as a development of the thought already present in the latter work that people commonly lack the leisure and training to make out their duties by the light of nature (together, perhaps, with a chastened appreciation that ethics might have some of the *difficulty* of mathematics).¹²³ It was, after all, in those same years that he completed the argument in the *Essay* assigning an inferior epistemological status to revelation, whether traditional or immediate. Finally, he did not, like Ockham and Hobbes, see God's authority as stemming from his omnipotence alone, but took considerable trouble to distinguish two aspects of that authority, the right to command and the power to punish. What gives rise to the former is not sheer omnipotence, but the peculiar indebtedness of the creature to its creator and conservator, *Ius* or right, Locke had argued in the *Essays*, is what distinguishes the legitimate authority or *potestas* of a king over his subjects from the mere *potentia* or power of a pirate to command his victims.¹²⁴ The distinction retained its full force in the *Essay*. What, then, prevented Locke from weakening still further the voluntarism in his theory and moving to a more purely rationalistic conception of the moral law as capable of being made out without explicit reference to its divine sanction?

Possible answers present themselves even without that detailed assessment of the sources and development of Locke's theological and ethical thought which is beyond the scope of the present work. It might, for example, be suggested that a more Thomistic view than Locke's would imply the existence of moral universals in the divine intellect accessible to human reason. That is indeed the model employed by Cudworth, who constructed an argument for God's existence from the possibility of moral knowledge. Obviously such Platonism would have been anathema to Locke, but he would seem already to have had the means of avoiding it in his general theory that universal

knowledge is concerned with abstract ideas rather than eternal archetypes. Moral universals, like mathematical ones, are mere mental abstractions. If it had any interesting rationale at all, the role of God's will in Locke's theory must be explicable as something other than a way of avoiding Platonist or Thomist ontology.

A more promising suggestion is that Locke took attempts to ground ethics on human nature or social utility alone to be incapable of explaining—or of supplying—what Mill later referred to as the 'binding force' of moral principles. Here his theory of value and his theory of motivation came together. He could not accept a view of ethics as a purely theoretical or intellectual inquiry into what is meet or fitting or in accordance with human excellence because it would then neither appear as unconditionally obligatory nor have any practical effect. In other words, his addiction to a voluntarist framework as an explicit element in moral reasoning was inextricably tied up with his hedonism, the role he ascribed to pleasure and pain (the agent's own pleasure and pain) as both the sole sources of our concept of value and the sole motives of action. A more purely intellectualist ethics demands a different theory of motivation and binding force: reason itself must motivate and oblige. As Clarke put it, perhaps with Locke in mind, whatever may be commanded, the 'original Obligation of all...is the eternal Reason of Things'.¹²⁵

A direct appeal to self-preservation and happiness in this life, on the other hand, while consonant with Locke's theory of motivation, could not, he evidently felt, actually motivate consistent and unconditional adherence even to the rules which uphold society, quite apart from its failure ultimately to explain our sense of obligation to obey them.¹²⁶ Crudely, it seems possible that the chief reason why Locke clung to a theocentric account of moral knowledge was that he saw no other way of explaining such knowledge compatibly with his concept-empiricism, or, indeed, of explaining it as practical knowledge.

Before an attempt, in the next chapter, to assess this at least respectably philosophical and perhaps disquieting motive, the present cursory account of the foundations of Locke's ethics should be completed by some discussion, first, of his psychological hedonism and, second, of his explanation of freedom, an argument of which he was himself particularly proud. The theory in fact underwent considerable revision between early editions of the *Essay*. In the first edition his doctrine was that we are motivated by 'happiness', or, more precisely, that 'the [*scilicet* appearance of] *greater Good is that alone which determines the Will*'.¹²⁷ But by the second edition what is now generally known as the problem of 'weakness of will' had stimulated a refinement in keeping with a stronger form of the principle that reason cannot move to action. It is a feeling or passion, namely some form of present uneasiness, rather than an intellectual or 'speculative' view of future happiness, which serves as the motive of any action:

Let a Man be never so well perswaded of the advantages of virtue, that it is as necessary to a Man, who has any great aims in this World, or hopes in the next, as food to life: yet till he *hungers and thirsts after righteousness*, till he feels an *uneasiness* in the want of it, his *will* will not be determin'd to any action in pursuit of this confessed greater good; but any other *uneasinesses* he feels in himself, shall...carry his *will* to other actions.¹²⁸

Indeed, Locke asked, how could *future* pleasure act on the will now? Uneasiness 'alone is present, and 'tis against the nature of things, that what is absent should operate, where it is not'.¹²⁹ Locke had come to rest with an extreme view which has been enormously important in psychological ethics, helping to shape even such an opposing theory as Kant's. That is the doctrine, not simply that all men pursue happiness in all their actions, but that only 'passions' and 'feelings', 'modes of pleasure and pain', are the immediate causes of action. Beliefs, even evaluative beliefs, are inert.

The problem of accounting for weakness of will is a traditional problem in ethics which stemmed originally from the Socratic (and highly 'intellectualist') paradox that no man does wrong knowingly. It was discussed at length by Aristotle. But what particularly troubled Locke was that even 'lively representations' of the afterlife, 'which they acknowledge both possible and probable too', can fail to turn people from sin. The problem is exacerbated by a consideration expressed in Pascal's famous 'wager'. The principle that 'the greater good in view' determines the will must be interpreted, Locke took it, in terms of two independently variable factors, the extent of the good and its probability as a consequence of action. But if a 'good in view' is infinitely greater than any other, as the joys of heaven are greater than earthly joys, then recognition of its bare possibility should outweigh the attractions of all rivals.¹³⁰ Yet it often fails to do so, and Locke's explanation is that the confessed greater good does not in such cases make a part of the agent's present happiness, i.e. its absence causes no uneasiness. People are commonly too occupied in satisfying their natural desires (and the equally solicitous itches due to bad habits) to be much concerned about anything else: 'the present is apt to carry it'.¹³¹ Locke of course regarded this situation as deplorable and contrary to reason. The remedy lies in the power people have to '*suspend* their desires, and stop them from determining their *wills* to any action' in order to examine the possibilities carefully and to reflect 'whether it be really of a nature in it self and consequences to make him happy, or no'.¹³² Somehow, Locke supposed, deliberation can not only increase information about means, but can achieve a correspondence between judgement and passion, and 'suit the relish of our Minds to the true intrinsick good or ill, that is in things'.¹³³ Reason, it seems, can at least work indirectly on the will by raising uneasiness where it is appropriate. Consequently, 'Morality, established upon its true Foundations, cannot but determine the Choice of any one that will but consider.' Failure to 'consider' is itself morally blameworthy, except when extreme pain or disturbance 'allows us not the liberty of thought'.¹³⁴

This incoherent, yet strangely attractive account of rational agency was set out in the context of a discussion of freewill. Locke, like St Augustine, Descartes and many others before and since his time, adopted the 'self-determinist' solution to that traditional problem. That a choice is determined by the agent's 'own desire guided by his own Judgement' does not detract from his liberty. On the contrary, Locke argued, that is what it is for a choice to be free, and those who postulate causal 'indeterminacy', whether it is supposed to occur before or after judgement, postulate a factor antithetical to liberty. One way or the other they turn choice into something unguided by judgement. Yet the simple picture employed in this argument of a chain of events (roughly, *deliberation* → *last judgement of good and evil* → *last uneasiness or desire* → *choice* → *action*) is complicated

by deliberation's being itself something which can be undertaken or not. Two powers come into the question, not one. There is 'the power to act or not to act according as the Mind directs', which seems to have been Locke's official definition of 'liberty'. But there is also the power to suspend the action of a particular desire in order to consider with due care whether the apparent good which gave rise to that desire is the real good for us, making 'a part of our real Happiness'.¹³⁵ Locke often suggested that it is this second power which really constitutes the liberty of rational agents, and which manifests the essential link between reason and freedom:

The first therefore and great use of Liberty, is to hinder blind Precipitancy; the principal exercise of Freedom is to stand still, open the eyes, look about, and take a view of the consequence of what we are going to do, as much as the weight of the matter requires.¹³⁶

In his account of this second power, Locke had apparently forgotten the rhetorical question which he asked about the first: 'For how can we think any one freer than to have the power to do what he will?'¹³⁷ Yet, not to speak further of the unresolved duality in his explanation of liberty, the second liberty or power he postulated would seem to give rise to an awkward question—awkward, at least, in the context and terms of his own argument. If everyone has the power to stop and think in the face of urgent desire, what determines that some people choose to exercise this power while others do not? What about those who perceive the advantages of pausing to deliberate, to allow feeling to match considered judgement, but who still fail to do so? If reason cannot directly move to action, how can it directly move to deliberation?

The only material Locke seems to have supplied for an answer to these questions is distinctly insensitive to the real problem:

Whatever necessity determines to the pursuit of real Bliss, the same necessity, with the same force establishes *suspence, deliberation*, and scrutiny of each successive desire, whether the satisfaction of it, does not interfere with our true happiness, and mislead us from it.

The passage is part of a sermon on the virtues of careful deliberation, rather than a response to objections to his account of liberty. The term 'necessity' might have been expressly chosen to fudge the distinction, set up by his own psychology, between moral 'obligation' and 'motive': between what is rationally or speculatively compelling, and what is so psychologically or practically:

For the inclination, and tendency of their nature to happiness is an obligation, and motive to them, to take care not to mistake or miss it; and so necessarily puts them upon caution, deliberation and wariness.¹³⁸

The problem, threatening regress or worse, of why some people in fact fail to pause

sufficiently to match motive with perceived obligation or self-interest is here passed by, although it was the existence of such people which led Locke to make explicit the distinction between 'judgement' and 'motive' in the first place. The capacity to reflect and deliberate is no doubt a condition of freedom of action and moral responsibility, but Locke did not succeed in giving a satisfactory account of that capacity in the terms of his emotive theory of motivation.

16

Reflections on Locke's ethics

Perhaps the most urgent philosophical question raised by voluntarist ethics is what to put in its place. That may seem an inappropriate judgement on a theory so long gone, but it is implied by the radical modern argument that the whole notion of ethics as a branch of objective knowledge, a notion firmly embedded in ordinary language, rests on assumptions no longer tenable, such as the assumption of a world ordered by its creator's intentions. According to this sceptical argument (which has been around at least since Nietzsche) the objective form of everyday moral reasoning and dispute is a relic, ripe for deconstruction, of a system of beliefs different from our own. As one commentator has pointed out, the argument is in effect the same as that advanced by Locke with considerable passion and emphasis, but in reverse. Unless right and wrong are the creation of God, we are left with principles of action invented by man and therefore utterly lacking the unconditional authority of the moral law as we ordinarily conceive of it, or have hitherto conceived of it.¹³⁹

More fundamental than the question which horn of this dilemma to embrace, of course, is the question whether the argument which sets it up is sound. Here it is relevant that the philosophical weak point of full-blooded voluntarism had been evident to its rationalist critics centuries before theism ceased to be in full health, through the recognition of the autonomy of moral judgement. Briefly, divine commands cannot be the source of moral distinctions since they and their author are themselves possible objects of moral evaluation. That is the reason for the traditional anti-voluntarist proposal that God is 'obliged' to act justly by his own knowledge and wisdom, a proposal designed to avoid both the paradox that the distinction between right and wrong is arbitrary (or that the judgement that God is good is a mere tautology) and the theologically unsavoury conclusion that God is constrained by something outside himself. Locke appears to have been unconcerned with that issue, happily ascribing justice and goodness to God in their plain sense. Such *insouciance* is understandable if it is right that his voluntarism was sustained less by the theological preferences usually associated with it, of which there is little sign in his writings, than by his felt need to explain our human sense of the authority and obligation of morality—not to speak of its practical effect on beings for whom pleasure and pain constitute both the only motive and the only source of value. Yet Locke's voluntarism suffered from effectively the same flaw as the full-blooded theory, for his argument required him to step outside his own explanation even of human obligation. If the obligation to obey the divine lawgiver is itself a moral obligation (in virtue of a *ius creationis*), then it cannot be the case that the idea of moral obligation is explicable in terms of the constraints of divine law. No law can create a duty to obey it by including that duty in its own content.

What is formally wrong with Locke's theory is not that he makes all moral obligation rest on one fundamental intuitive obligation to act in accordance with the will or law of our Creator: it is rather that he combined that doctrine with an analysis of the idea of obligation which would make it impossible for any obligation so to transcend all law. The incoherence apparently went even deeper than that, since our intuitive obligation to God is sometimes presented as just one, if the strongest and most certain, of a number of parallel obligations. These comprise not only the duty of a son to his father but the rights of any craftsman over his production, the source or core of moral rights to property.¹⁴⁰ To allow the existence of such independent moral truths is of course incompatible with the main thesis. But if we take the internal inconsistency of Locke's theory to be an instantiation of the necessary incoherence of any attempt to found moral obligation on a divine order, it will perhaps be less tempting to suppose that the notion of objective moral truth owes its existence to an outdated belief in such an order. That is not to say that the 'binding force' of morals does not call for explanation. It is just that the theistic explanation never worked anyway, even in its own terms, so that its passing is hardly a matter of regret for the moral objectivist.

Nevertheless the structure of Locke's failure may repay further study. First, his explanation of the meaning of the words 'good' and 'evil' in terms of pleasure and pain leaves him open to the charge of 'naturalism', i.e. of ignoring the semantic difference between description and evaluation. In the context of his project of tracing our concepts to experience, however, his argument has considerable force. The question of how experience gives us ideas of good and evil was for him the question of what it is in our experience of things which makes anything matter to us, and which makes one thing preferable to another. Without something which constitutes a ground for preference there could be no evaluative concepts. The hedonist holds, not implausibly, that the only ingredients of experience which supply such grounds are pleasure and pain, happiness and misery. He may express his claim, as Locke did, by defining 'good' and 'evil' in terms of pleasure and pain (as Bentham also did, for whom, when interpreted with reference to his hedonistic principle of utility, 'the words *ought*, and *right* and *wrong* and others of that stamp, have a meaning: when otherwise, they have none').¹⁴¹ But the claim itself is not just bad semantics, to be refuted by G.E. Moore's famous argument that, as far as the meaning of 'good' is concerned, anything whatsoever may be good. Moore's principle is true simply because, as a dictionary may tell us, 'good' is the most general term of commendation. It is not the *meaning* of 'good', but something else, which makes it merely unintelligible to claim that stones or blue things or bare feet are good in themselves, to be pursued for their own sake. Yet without an intelligible basis for preference we would have no concept of the preferable. Nothing would matter (no doubt, *per impossibile*), and there would be neither commendation nor words of commendation. In that way, at least, it is tenable that hedonism explains how the words 'good' and 'bad' have meaning. That is not to deny that Locke, on a strict analysis, sinned against Moore's semantic principle; it is rather to say that the sin is venial. It is a more profound criticism of his doctrine to argue (as Moore did go on to argue, if much less systematically and impressively than Kant) that there are some considerations which make sense as reasons for action and preference, but which are not in any way hedonistic.¹⁴²

It was of the essence of Locke's hedonism that it had two faces, as an empiricist explanation of our concepts of value and as a theory of motivation. The only things that matter are the only things that motivate. As Bentham later put his similar view: 'Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do.'¹⁴³ It is a condition of the coherence of such a theory that its two faces do not come apart. Yet we have already seen that that is just what did happen in the course of Locke's account of obligation, since his distinction between right and power, and his recognition of a moral obligation to uphold the command of our Creator (an obligation from which Locke derived other duties and human rights, such as the right to punish transgression of the law) constituted a recognition of other values than flow from pleasure and pain. His emphasis on God's power to punish and reward as what supplies the motive for obedience is in effect an expression of his assumption that the mere recognition of right, whether God's right or other people's, cannot by itself move to action. Yet such a combination of psychological egoism with a conception of rationally intelligible rights and duties takes us away from what made psychological egoism plausible in the first place, i.e. the thought that pleasure and pain supply the only intelligible reasons or grounds for action. For to recognize any moral right or claim of others (whether or not that right is supposed based on human interests hedonistically defined) is to recognize a reason or (in an ordinary sense) a 'motive' for action, whatever the consequences for oneself of observing or contravening the right on the present occasion. So Locke had in effect got himself into the curious position of recognizing a reason which he held not to be (in his sense) a motive, i.e. a reason which will not actually move us to action. What moves us to obey the law cannot be God's right to obedience, but must be his power or, rather, our own fear and trembling.

Locke was in the grip of two related conceptions: first, that an effective reason for action exists only in so far as the action contributes, or is taken to contribute, to the agent's own happiness; and, second, (as our failure always to pursue the perceived good was supposed to demonstrate) that rational judgement does not motivate in and by itself. Even without his theistic framework the two conceptions are commonly linked, as in the popular egoist argument that apparently altruistic actions are performed in order to allay the uneasiness we feel at the suffering of others. To consider the second conception first, the notion of uneasiness supplied Locke with a poor solution to the problem of weakness of will if only because in any particular case there may be no such emotional disturbance, or what there is may lie on the other side. Cold calculation of worldly profit may motivate an immoral action even while recognition of its consequences for others gives rise to considerable uneasiness. Here victorious reason is in the wrong, defeated passion in the right. Locke's other argument for the role of present uneasiness (i.e. that future pleasure, being absent, cannot act now) betrays how far he simply presupposed that judgement, even judgement about one's own interest, cannot move to action. If we drink when we are not thirsty, he assumed, then some other passion, a felt 'thirst' for health, intoxication or some other supposed good, must be present. It is doubtless true that, if we drink intentionally, we want or desire something. Yet what is problematic is the combination of this platitude with a sharp distinction between desire and the intellectual

acceptance of an action-guiding judgement. Such a distinction interprets desire as an emotion or 'passion' in something like Locke's sense, a passion which may or may not be raised by the rational judgement.

There are, of course, two opposed ways of bringing desire and judgement together: one is to allow that the notion of desire includes the rational acceptance of action-guiding judgements, and the other, adopted in this century by 'emotivism' and its offshoots, is to treat all action-guiding judgements as ultimately expressions of non-rational feeling. The former is preferable, but not without its dangers. For it is not enough to insist that judgement and reason can move to action in the absence of feeling, or in the face of feeling. The classic exponents of that view, such as Plato, Butler and Kant, characteristically thought in terms of a rigid opposition between the rational or ethical motive on the one hand and natural passions and inclinations on the other. Two sorts of desire are postulated, one rational and conscientious, the other passionate and blind. Yet as the example of moral uneasiness itself might suggest, such an opposition is itself open to question. Emotions themselves necessarily involve value-judgements, and to act from the motive of fear or anger (not to speak of a passion for justice) is, among other things, to act for a certain sort of reason. Feelings, like actions (although not similarly voluntary), may be appropriate, justified and even demanded by the circumstances. Indignation, sympathy, guilt, admiration and amusement may themselves be manifestations of rationality and intelligence. Sense and sensibility are less distinct than most philosophers have supposed. That is why, although some desires or wants are passionate and others are not, 'desire' and 'want' are not therefore ambiguous terms.

The mutually supportive link between emotivist theories of motivation, psychological egoism and ethical egoism has been explored in one of the most effective recent arguments against egoism. Its author, Thomas Nagel, argues that psychological egoism begins to lose its charm once the principle that only feelings move to action is called into doubt. That is done if we can show that even in purely self-concerned practical reasoning a wise course of actions may be preferred without present feeling or emotion. Once we see how unproblematic it is that we should be motivated in our investments or in ordering our dinner by the mere calculation of future benefit or harm, then the way is open for us to see how unproblematic it is that we should be motivated by the mere expectation of others' benefit or harm.¹⁴⁴ It is also, one might add, open to us to recognize that we can be rationally motivated by other considerations altogether than good or bad consequences hedonistically conceived.

Although Locke was a psychological egoist, he was not an ethical egoist, and it is possible to see his intense dedication to the theistic, voluntarist framework as the result of his felt need to reconcile the notion of an impartial, binding moral law with the notion that our actions are never impartial. Because they adopted broadly the same psychology, some later utilitarians faced the same unresolved problem. Mill, for example, defined the 'motive' of an action as the feeling which makes the agent will so to act. He also held, notoriously, 'not only that people desire [their own] happiness, but that they never desire anything else' and that 'to desire anything, except in proportion as the idea of it is pleasant, is a physical and metaphysical impossibility'. Consequently, although he spared himself the embarrassment of an intelligible duty to the creator, he was saddled with a

task, very like Locke's, of fitting together psychological egoism with ethical altruism or impartiality. He tried to solve the problem of the 'sanction' or 'binding force' of morality, as we might expect, in terms of special motivating feelings, 'conscientious' feelings produced by training and authority and capable of being attached to any system of ethics whatsoever. The system to which they are most naturally, readily and permanently attached, however, is precisely the system founded on those 'social feelings' which make our happiness depend on the happiness of others and which, he argued, can and should be endorsed and cultivated by education. In all this, however, the fundamental difficulty remained untouched. For if, in addition to impartial ethical reasons, we needed sympathetic 'feelings' or 'motives' to lead us to act on others' behalf, then ethical reasons could hardly be expected to lead us to prefer and encourage social feelings over other, more self-centred emotions. We might encourage such feelings in other people out of self-interest, and conform to a social code ourselves for the same reason, but that takes us back to Locke's far from morally binding 'law of opinion'. The problem is that Mill's theory of motivation, just as much as Locke's, drives a wedge between judgement and action which makes nonsense of the whole notion of moral reasoning as a branch of practical reasoning.¹⁴⁵

How, then, might our sense of the 'binding force' of morality be explained? We need, it seems, to invoke a combination of two factors, neither of which is adequate by itself. In part, the moral judgement of obligation appears 'binding' in so far as it is rational, and appeals to intelligible ends, rights or ideals. Yet the pure rationality of the judgement cannot by itself be the full explanation of its binding force just because the rational pursuit of one's own interest, while unconditionally rational, is morally neutral. As far as concerns what is strictly in my own interest, I am under no obligations. I have, as it were, a perfect right to act irrationally. 'Binding' practical principles, on the other hand, arise precisely when the interests and rights of others constitute intelligible grounds for me to act. The demands made on me by the moral law are in one way or another the reasonable demands made on me by others—and by 'demands', of course, is meant not only voiced demands. There is no more transcendental foundation of moral obligation, or of the analogy with law, than that.¹⁴⁶ On the other hand, obligation is no more created by us than by God. To suppose that the 'death of God' removed the basis of obligation is at best a confused way of expressing scepticism about the rationality of such demands or claims, a sort of moral scepticism which is hardly to be confronted here. Yet one concession to the sceptic is easy to make. A reasonable ethics is likely to be considerably less rigorous and systematic than Locke's analogies with mathematics and law would suggest.

Part III
Identity

Introduction to Part III

The chapter on identity and diversity added to the second edition, apparently at the instigation of Molyneux, stands in some respects as a natural culmination of Locke's system, if in others as an awkward afterthought. Like very many seventeenth-century discussions of the *principium individuationis*, its real concern is with the possibility of immortality. Yet it constitutes one of the most philosophically interesting chapters of the *Essay* for the modern reader, as well as, historically speaking, one of the most controversial and influential. Locke's anti-dogmatic programme, combined with his 'reasonable' Christianity, required him to demonstrate that scepticism about the essence of that which thinks in us is compatible with a belief in immortality: 'All the great Ends of Morality and Religion, are well enough secured, without philosophical Proofs of the Soul's Immateriality.'¹ In effect he argued that consciousness of oneself as a thinking thing is not, as the Cartesians, Leibniz and others supposed, immediate awareness of the simple essence of an immaterial substance. For the perceived unity of the thinking thing (including its continuity over time) just *is* the given unity of consciousness essentially involved in the reflexivity of thought. This passage (in particular, II.xxvii.9) was no doubt a crucial source of Kant's conception of the unity of apperception as a necessary condition of experience, but all Locke meant to say was that the perceived unity of consciousness is no more than a unity at the phenomenal level, and that (as Kant agreed) no conclusions can be drawn from it as to the ultimate unity or simplicity of the underlying substance. As such, he held, it is both necessary and sufficient for the purposes of a satisfactory account of moral obligation as obligation to a divine lawgiver with the power of eternal reward and punishment.

The argument was set within a nominalist or conceptualist, anti-Aristotelian approach to individuation which such earlier writers as Hobbes, Spinoza and Boyle had begun to explore in their different ways. Locke's attempt to present this approach in a form adequate for his eschatological purposes (immortality, after all, can hardly be a mere matter of words or concepts) led him into a theory both teasingly indeterminate in itself and at odds with some of the presuppositions of the first edition onslaught on Aristotelian species. Many alleged contradictions in Locke's thought melt away on closer examination, but in this case the charge sticks. In the first edition he was prepared to assume that an individual could in principle survive transmutation from horse into (say) sheep or swan, and indeed that I might continue to exist as an individual after the loss, not only of an 'accident' such as a characteristic colour, nor even of the shape which makes me human, but also of reason, memory, sense, understanding and life itself.² Yet in the second edition, he came also to propose that life *constitutes* the individual animal, while reason and memory, as 'consciousness', constitute the self. Nevertheless the

inconsistency is understandable, if not dissoluble. For each of these incompatible positions had its source in argument directed against Aristotelian forms. The two lines of argument in question could coexist (and did so coexist in a single book by Boyle) in forms in which they were compatible and even complementary. Locke in effect developed one of them in the interests of eschatology to a point at which its claims contradict those of the other line of argument. This process, and some of its philosophical implications, will be examined below.

Boyle's argument had not been concerned with immortality, but with substantial change and, paradigmatically, with death. How are we to understand biological death, and the being of the individual that dies? It is Locke's account of the significance of biological life and death, supplied as a necessary preliminary to his explanation of personal survival, which most strongly indicates a connection between his general theory of identity and earlier anti-Aristotelian nominalism. His treatment of biological death had, indeed, a motive independent of his interest in the afterlife. For he might well have felt called upon further to define his position in relation to the Aristotelian doctrine (or rather, perhaps, in relation to the natural intuition which can seem to give significant support to that doctrine) that at death the individual goes out of existence. In the first edition he explicitly rejected any such intuition, but in the second he came to make important concessions to it. It will be suggested that in this case first thoughts were better than second. They were certainly more straightforward, and the sheer indeterminacy of Locke's later theory makes it more difficult than usual, and probably less desirable, to separate exposition from philosophical criticism.

Locke on 'masses of matter'

The Aristotelian understanding of death is a piece of outdated scientific ontology, a function of the Aristotelian theory of generation, growth, nourishment and change in general. Aristotle held that a natural stuff (the specific form of which is always itself reducible to a certain combination of the fundamental 'opposites', hot and cold, wet and dry) can have the active capacity to convert other matter into its own kind, as if a small amount of water added to a large amount of wine really became wine. Alchemists later accepted that metals grow in this way in the earth and in the fire. Fire itself was taken to be a natural substance with an active capacity to transform the inflammable. A fire fed with logs grows. The flesh or tissue of the various parts of the body was supposed to grow in the same way, transforming the matter of food so that new particles of tissue come into existence. At the same time other particles of tissue cease to be. Aristotle himself used the somewhat Heraclitean metaphor for growth and nourishment of 'flowing water that is measured by one and the same measure', 'some flowing out and some flowing in'. The vessel or 'measure' represents the active form of living flesh, or, indeed, the form of the animal. The water in the vessel represents the 'matter' of the tissue, previously the matter of the food. This matter, considered as surviving the change of substantial form or nature, must be regarded as indeterminate even in quantity: i.e. as '*materia prima*'. The form of this or that sort of tissue is subordinate to, perhaps an aspect of, the form of the living individual. Death is the going out of existence of the form of the animal, hence of the animal itself. The active forms of flesh, bone and so forth go out of existence *ipso facto*: dead 'flesh' is inactive as such, and therefore not flesh, however much it may look like it. It is as if the vessel had suddenly been annihilated, while the matter, which was a moment before contained in it, temporarily and imperfectly kept its shape. Thus Aristotle accepted the Heraclitean flux of 'matter', but death, and substantial change generally, involves an absolute 'ceasing to be' of the substance.³

The corpuscularians, in rejecting substantial forms, also denied that substantial generation or corruption can occur naturally, without an extraordinary act on God's part of creation or annihilation. They therefore tended to regard everyday talk of things' coming or ceasing to exist as a *façon de parler*, relative to our classification or naming of things. Thus Boyle argued that a body (i.e. a kind of stuff)

is said to be generated, when it first appears clothed with all those qualities upon whose account men have been pleased to call some bodies *stones*: others, *metals*; others, *salts*, etc.... [And] when a body comes to lose all or any of those accidents that are essential and necessary to the constituting of such a body, it is

then said to be corrupted or destroyed, and is no more a body of *that kind*, but loses its title to its former denomination.

The 'body' still exists as a body, but 'tis no longer *such a body*, as 'twas before, but perisheth in the capacity of a body of that kind'.⁴ Death received comparable treatment: a dead animal is like a watermill, prevented from working in a drought or frost, which has ceased to perform the function in virtue of which it is called a mill. The difference is that the animal or plant begins to decompose so that it will not work again, but in neither case does anything substantial perish. Moreover, Boyle noted, a stick or piece of aloe may sometimes revive in the right conditions even years after being cut, like a mill which resumes working with the renewed motion of water. But, in any case, even when life has ceased for ever, other functions of the individual may continue: 'the Body may upon the account of the more permanent structure of its stabler parts retain a fitness for divers of the same purposes it served before'.⁵ There is nothing special about the function of life in itself, and we can continue to talk of the same individual's being in existence as long as there exists a stable structure.

Locke evidently felt the need to present his version of the mechanist treatment of individual substances within a more ambitious theory of the identity of individuals of any category whatsoever. First, particulars *A* and *B* are numerically diverse if they exist, at one and the same time, at different places; or if they exist at the same time and place, but are, in a certain unexplained sense, of different 'kinds'. If *A* and *B* are of the same kind and exist at one time at the same place, then *A* is identical with *B*. But some particulars can exist through time, so that if *A* is known to have existed at *t*, and *B* at *t+n*, it is still possible that *A* is identical with *B*. In such a case a necessary condition for *A*'s being identical with *B* is that *A* should be of the same kind as *B*, and that there should exist an individual of that kind at all times between *t* and *t+n*. The 'kind' of anything, it seems, is determined by an 'idea': a being is 'of the same kind' as long as it corresponds to the same idea. In other words, this particular red quality, this thing's redness, remains a particular 'of the same kind', and hence *exists*, only as long as it corresponds to the same idea, i.e. as long as the predicate 'red' is satisfied. As Locke put it (placing 'ideas' in things themselves), 'in this consists *Identity*, when the *Ideas* it is attributed to vary not at all from what they were at that moment, wherein we consider their former existence, and to which we compare the present'.⁶ So expressed, this principle might seem to make it impossible for any particular to undergo alteration at all, but that, of course, was not Locke's intention. In the case of substances, at any rate, the 'idea' in question was supposed not to rule out alteration, but to set the limits within which alteration is possible.

The satisfaction of a predicate over time, without intermission, was not, of course, supposed to be enough for the continuity of a particular. Continuity through space was seen to be important, and Locke accordingly turned, in II.xxvii.2, to the special role of substances as the primary occupants of space. We have ideas of three general sorts of substance: God, finite intelligences and bodies. No problem of identity can arise about God, since God is omnipresent, eternal and unalterable. A particular 'finite intelligence', however, like everything else finite,⁷ uniquely (of its kind) enjoys a particular 'determinate time and place of beginning to exist'. Its identity at later times is determined

by its continuous relation through space over time to that origin, i.e. by what is now called 'spatio-temporal continuity'.⁸ Locke did not explain what it is for a finite intelligence to occupy a place, which he elsewhere found a difficult but necessary conception.⁹ But that problem does not arise for 'bodies' or 'parcels of matter', to which, he said, exactly the same principle applies. It is the perspicuousness (or at least relative intelligibility) of the place-occupancy and spatio-temporal continuity of substances which supplies what is missing in the general account of identity. The individuation of substances is primary: 'All other things being but Modes or Relations ultimately terminated in Substances, the Identity and Diversity of each particular Existence of them too will be by the same way determined.'¹⁰

The 'much enquired after' *principium individuationis*, Locke concluded, is existence itself, 'which determines a Being of any sort to a particular time and place incommunicable to two Beings of the same kind'. The medieval doctrine that the principle of individuation of particular things is existence seems to have been a response to the supposed problem of how beings of the same nature are distinguished from one another: since particular existence is prior to the existence of universal natures, no special explanation of the individuality of particulars is called for (no *haecceitas* or 'thisness') beyond that existence itself.¹¹ Locke no doubt agreed with that, but the point of *his* argument is not only that the existence in question is existence in space and time, but that it is the existence of a thing of a determinate kind. Consequently, even if the metaphysical *principium individuationis* is 'existence itself, what matters with respect to questions of identity is that some 'kind' is instantiated, some general predicate satisfied, at determinate times and places. The meaning of the predicate, the idea with which it is associated, determines the criterion of identity of the particular. All this may seem to constitute a complete, if exceedingly general theory of the individuation and identity over time of particulars. Locke now turned to the particulars which interested him, biological individuals and rational agents.

Locke took it that his general account is unproblematic with respect to non-substances and simple substances. By a 'simple' substance he meant an ultimate particle or atom which, by hypothesis, is unchanging.¹² It is with compound substances, which can undergo change, that he felt the need for more argument. First, he took it that there is a substantial object, corresponding to the idea of a united 'mass' of atoms or compound 'body', the identity of which is determined simply by the identity of its elements and their conjunction: 'if one of these Atoms be taken away, or one new one added, it is no longer the same Mass, or the same Body'. The ideas determining the identity of biological individuals, however, are different. In the growth and nourishment of such things there is allowed to be 'a manifest change of the parts'. Consequently, if we point to a 'plant' at one time and to 'a great tree' at a later time, we point at different 'masses of matter' but may be pointing at the same oak: here '*Identity* is not applied to the same thing.' A 'mass of matter' is 'only the Cohesion of Particles of Matter any how united', while an oak is 'such a disposition of them as constitutes the parts of an Oak; and such an Organization of those parts' as fulfils the functions of nourishment and growth, 'in which consists the vegetable Life'. The continuity of a plant's existence is the continuity of the one common life which 'vitally unites' its parts. Like Boyle, Locke drew the comparison with the

functioning of a machine, making the point that a machine differs only in that it depends for its motion on an independent force, and can start again after it has stopped.¹³

This whole argument raises a number of fundamental questions, but we might as well start by considering Locke's curious and artificial notion of a 'mass of matter'. It is of course true that, in some sense of 'individual', we can identify two substantial, material individuals of different kinds or categories occupying, within the spatial outlines of a horse, the same place at the same time. These are the horse itself and the quantity of matter currently composing it. Yet Locke's 'mass of matter' is evidently not a *quantity* of matter, since it is essentially 'united' physically. A quantity, on the other hand, is an entity analogous to a plurality or purely ideal aggregate. To trace such an entity through time is to trace its members or parts, and its history is simply the sum of their histories. For example, it is a mere convenience, with respect to its ostensive identification, if a particular number of people is present at a meeting. That particular number of people can continue to be identified however widely it is dispersed, and it survives just as long as every person in question survives. Similarly, that a particular quantity of stuff currently constitutes a horse is a temporary deictic convenience, temporarily facilitating (as well, most probably, as stimulating) the identification of just *this* stuff through such expressions as 'the present matter of this horse'. The quantity of matter is not as such a naturally bounded or united object. It might have existed without ever having been physically united, like the quantity of gold denoted by the expression 'the gold within the boundaries of Oxford at noon, 10 July 1980'. Its survival or continuity is natural and independent of our concepts just because its survival is no more nor less than the survival of every ultimate particle (or whatever) that makes it up. But its unity is notional or ideal. Hence the conception of its *gaining* a particle, or of its losing one except by the particular's destruction, would be meaningless. Precisely for that reason we call it a 'quantity' of gold.

There is nothing mysterious about the route by which Locke arrived at his idiosyncratic conception of a 'mass of matter'. He attempted directly to derive from his consideration of 'simple' individual substances a perspicuous principle of individuation for a complex, coherent, thing-like, material entity or 'body', but a principle which is left uncomplicated by the factor of 'organization', and so can be contrasted in this respect with the principle of individuation of living creatures. It is more than plausible that a stimulus for his argument here was Hobbes' wellknown discussion of the 'great controversy among philosophers', whether 'individuity' should be placed in 'the unity of matter' or 'the unity of form'. Either answer, Hobbes had claimed, if made quite general, leads to paradox. To place individuality in the unity of matter would entail that no man was the same for more than a moment, so that 'he that sins, and he that is punished, should not be the same man'. To place it in unity of form would allow that 'two bodies existing both at once, would be one and the same numerical body'. This second paradox is supposed to follow from the story of the ship of Theseus, which was repaired plank by plank until none of the original matter was left. The old planks were preserved and eventually made into a ship again, so that 'this, without doubt, had also been the same numerical ship with that which was at the beginning'.

Hobbes' paradoxical conclusion might, of course, be disputed in its own terms, since it

might be claimed that the ship made out of the old planks has not existed continuously, and that continuous existence is a condition of identity. (Against this, it might be said that a ship can exist in pieces.) But it is Hobbes' solution to the paradox which is more to the present point:

But we must consider by what name anything is called, when we inquire concerning the *identity* of it. For it is one thing to ask concerning Socrates, whether he be the same man, and another to ask whether he be the same body; for his body, when he is old, cannot be the same it was when he was an infant, by reason of the difference of magnitude; for one body has always one and the same magnitude; yet, nevertheless, he may be the same man. And therefore, whensoever the name, by which it is asked whether a thing be the same as it was, is given it for the matter only, then, if the matter be the same, the thing also is *individually* the same; as the water, which was in the sea, is the same which is afterwards in the cloud; and any body is the same, whether the parts of it be put together or dispersed; or whether it be congealed, or dissolved.¹⁴

The name 'man', Hobbes' argument continued, is given 'for such form as is the beginning of motion', so that 'that man will always be the same, whose actions and thoughts proceed all from the same beginning of motion, namely, that which was in his generation'. The name 'ship' Hobbes claimed, is given neither for a 'form' or source of action nor simply for matter, but for an accident of shape. It 'signifies matter so figured'. It is therefore necessary for identity that both the shape and the matter are unchanged, the latter because the accident could not remain numerically the same unless its subject remained the same. If the matter is partly replaced, he adds, 'then the ship will be partly the same, and partly not the same'.

Locke did not follow Hobbes closely in all respects but, like Hobbes, he believed that the identity of some individual objects involves the identity of their matter, and, like Hobbes, he operated with an ambivalent notion of a *body*. Hobbes accounted for the 'body' of Socrates as if it were simply the stuff composing Socrates at some time (consider, for example, the principle that its quantity cannot change). Yet, despite the point that dispersal is irrelevant to the identity of a body in this sense, he also treats it as the discrete and unitary object which we would normally mean by the expression 'the body of Socrates'. Locke was more explicit over unity, for his argument more obviously requires that the body be a united thing. But he crucially neglected the point that 'organization', in the form of a continuing dynamic relationship between its parts, is a condition of the physical or material unity which he ascribed to the 'mass' or 'body'. Without material coherence we can make no sense of the notion of an atom's leaving or joining the 'mass'. Yet as soon as we introduce physical coherence into our conception of the 'mass' at any one time, we have taken account of the unity of that which can *survive* the loss or addition of particles. We cannot consistently employ material coherence as a principle of unity and at the same time reject it as a principle of continuity. Unless all reference to 'unity' is dropped from the account of the 'mass' (in which case we are left simply with the *quantity of stuff*), we shall find that we are talking about something like a

lump or chunk or unitary *mass of stuff* in an ordinary, non-technical sense, i.e. an entity which exists as long as it is coherent, and which can survive accretion and diminution according to the circumstances and the natural basis of its existence as a coherent, discrete thing. The expression 'the same mass of atoms' is thus not, as Locke supposed, equivalent to the expression 'a mass of the same atoms'. In this ordinary sense of 'mass', it is not possible to distinguish the coherent 'mass of matter' from the horse or plant occupying the same place at the same time. For the natural basis of the unity or coherence of the 'mass of matter' is not distinct from the natural basis of the unity of the horse or plant. It is true that this claim may be opposed on quite other grounds than those at present in question, most notably by the argument (to which I will return) that the horse ceases to exist at death and so is outlasted by the coherent mass of matter, or body, which was alive and is now dead. But the present point is simply that, given that there is a coherent material object which survives changes of matter (call it the horse), we cannot distinguish from that object a second united material object which fails to survive changes of matter.

It is worth considering precisely how Locke's 'mass of matter' manages to fall so ungracefully between two stools, being neither the quantity of matter nor the coherent material object. His attempt to conflate the two raises yet again the question how far entities can be 'constructed' by *fiat*, by the stipulation of 'principles of individuation', 'criteria of identity' or 'existence-conditions'. For, as we have seen, there are intuitively perspicuous constraints on such a procedure which philosophers commonly overlook. At the least, what is stipulated must be paid for: stipulation has its cost. We cannot, for example, independently stipulate *both* the existence-conditions *and* the category of an entity. As we have seen (and as Aristotle saw), we cannot 'construct' a substantial individual by treating the substance-plus-accident definition of 'baker', *one who bakes*, as the introduction of a new principle of individuation achieved by limiting the principle of individuation of a genuine substance. We cannot treat a *baker* as a thing which comes into existence in the school of cookery and ceases to exist on retirement. We may speak as if we can ('the poet Wordsworth died long before the man') but to take such talk seriously is a logical blunder. Logically speaking, the baker *is* the man, and what died before the man Wordsworth was his lyrical capacity, not a poet. But these intuitions are only *fully* supported and explained when we fully recognize that 'man' itself is not an arbitrarily formed predicate which sets the boundary-conditions of the individual in unbounded reality. It is, on the contrary, a general name given to independently identifiable, naturally bounded individuals in virtue of their belonging to a kind which is itself, as it happens, naturally bounded. Only when the implications of this point are recognized do we see why *man* is a 'sortal' concept while *baker* and *poet* are not. Or why 'red admiral' is, but 'caterpillar' is not, a sortal predicate.

The artificial, would-be principle of individuation of Locke's 'mass of matter' is arrived at, in effect, by just such an abortive attempt to limit an already intelligible principle of individuation. It can be approached by two routes. His 'mass' seems to be identifiable, on the one hand, as the quantity of matter composing an indicated unitary object *plus* the accident of its composing that object. On the other hand, it seems to be identifiable as an indicated unitary object *plus* the accident of its being composed of just

the same matter as composes it at the time at which it is indicated. It is self-evident that, if these composite notions or formulae really introduced new criteria of identity, they would converge and both individuate the same thing. Yet they do not do so. For this united-thing-composed-of-just-this-quantity-of-matter is a united thing, e.g. a horse, just as surely as a palfrey (a horse-trained-to-amble) is a horse. By the same token, this quantity-of-matter-just-as-long-as-it-constitutes-this-united-thing is a quantity of matter. Each 'individual' can survive the loss of the specified accident as readily as a baker can survive a change of occupation. There is no room for a third material or substantial individual between the two of which we normally speak and think. If there is an entity which ceases to exist whenever some of the matter of a horse is replaced, then it is not a substance but a 'mode' seldom directly mentioned in discourse but characterizable either as *this horse's being composed of this matter*, or as *this matter's composing this horse*. Such a mode is a relational situation or state of affairs which can be conceived of as an episode in the life-history either of the horse or of the matter. The individuality of this curious entity, unlike that of a united mass of matter, is ideal or notional, rather than natural or given.¹⁵

Our discussion so far yields two conclusions. First, Locke did not in II.xxvii.3 succeed in giving an example of two complex, united material things of different kinds occupying the same place at the same time; nor did he say anything in that paragraph which should lead us to suppose that such an example could ever be given. Second, there are intelligible constraints on the 'construction' of entities. We cannot expect to introduce a genuine 'principle of individuation' or 'criterion of identity' for a substance by stipulation or arbitrary definition. One might add that a satisfactory theory of the identity of particular things or substances is neither more nor less than an explanation of these constraints, their basis and their consequences.

Locke on living things

Let us now turn to the question whether an animal or plant is a thing which ceases to exist at death. That question is in effect the question whether ‘living-thing’ or ‘thing-which-is-alive’ is a ‘simple’ substance-term or a ‘compound’ term of the substance-plus-accident form. When the Aristotelian ontology of functional specific forms was accepted, it seemed clear what answer should be given. Now it may not be so clear.

The view that I have attributed to Boyle, although it does not represent everything he said on the subject, is simple but unsatisfying. Like Locke, he took it that a thing-sortal owes its meaning to a more or less arbitrary definition based on observation, i.e. to the selection of certain of the ways in which a physically unified thing acts as one. When all or, indeed, any of the functions or powers selected for the definition of the sort have fallen away, the thing has ceased to exist as a thing ‘of that kind’ or ‘in that capacity’. Boyle’s account was broadly Hobbesian or nominalist in so far as he would have attributed our saying of a broken watch that it has perished (i.e. in the capacity of a watch) to our possession of the noun ‘watch’ semantically linked to the function in question. His thesis was not so much serious ontology or logic, as the dismissive explanation of a supposed *façon de parler*. It could conveniently be extended to just those examples of ‘alteration’ which the Aristotelians would have contrasted with real substantial change or ceasing-to-be. Thus a horse could be said to be destroyed as a racehorse when crippled, or to cease to exist as a stallion when gelded, and so forth. In effect Boyle’s explanation reduces to a denial of the Aristotelian distinction between substantial and accidental change. He indeed explicitly held that the only absolute ceasing to be is the annihilation of substance itself, whether matter or spirit, through the intervention of God. As a result he did not focus very clearly on the disintegration of particular complex objects, which, to judge by his remarks about dead plants, he might seem to have been prepared to treat as the limiting case of a compound’s perishing in a certain capacity, by the loss of the last of its powers to act as one thing. Yet such final disintegration and dissociation of parts just is, surely, the absolute ceasing to be of a complex, material, united individual.¹⁶

Like Boyle, Locke in the first edition explicitly envisaged the biological individual’s existing after death, but he did not very explicitly propose that we can talk of its perishing relatively, as a thing of a certain kind. Still less did he seriously suggest that what remains in a way is, and in a way is not, what was alive. In the chapter on identity, however, it does occasionally seem that he had in mind that extreme conception of relative identity, associated today with the name of Peter Geach but apparently present in the passage from Hobbes quoted in the previous chapter, which would allow that individuals *A* and *B* could be the same individual under one denomination, but different individuals under another.

¹⁷ If such an interpretation could be put upon the general theory of identity of II.xxvii.1, the theory which links the identity of all particulars rigidly to membership of the same 'kind', then that theory could at least seem to be made compatible with the assumptions of the first edition that a particular substance can change kinds or species, and that one substantial individual can be paradigmatically associated with indefinitely many different specific or sortal ideas in the minds and systems of classification of different men. The logical difference between substances and modes could then be characterized by saying that a particular substance, unlike a particular mode, can be a member of many different species or 'sorts' at once; and that a particular mode, unlike a particular substance, is defined or bounded absolutely by any species of which it is a member. Yet unless such a model were offered, as Boyle can be taken to have offered it, simply in explanation of a certain figurative way of speaking about substances, it would be hard even to begin to take it seriously. For it leads directly to the paradox that one and the same particular thing has indefinitely many actual, distinct life-histories, each relative to a different sortal idea or nominal essence.¹⁸

Locke's general theory may, however, deserve to be understood rather differently. For he went out of his way to state that, from the principle connecting identity with 'the ideas it is attributed to' (i.e. with membership of a 'kind'), together with the principle that two things of the same kind cannot occupy the same place at the same time, we can derive the principle that one thing cannot have two beginnings, or two things one beginning in time and space. The conclusion may be no more than common sense, but the derivation would be impossible unless the first premise were taken to exclude the same thing's being a member of more than one 'kind'. In that case Locke was logically committed to rejecting Boylean relative coming and ceasing to be, or at least to taking such talk as unseriously as Boyle did himself. If he accepted that in cases of the type in question (such as death) something does properly cease to exist, then he was committed to accepting that that thing is absolutely distinct from what continues to exist (such as the body). It would follow that he was committed to holding that things of different kinds or sorts have been occupying the same place at the same time.

Yet if that were his view of these cases, if, that is to say, his principle that two things of *the same kind* cannot exist in the same place at the same time were intended to allow that physical objects of different species or sorts *can* do so, then Locke would be saying something not only alarming in its ontological consequences, but also incompatible with his previous conception of the relation between individual and species. For the possibility, so often alleged with emphasis elsewhere in the *Essay*, that the same particular can instantiate indefinitely many species at once by satisfying indefinitely many abstract ideas or nominal essences would need radical reinterpretation, i.e. as the possibility that indefinitely many complex material particulars should occupy the same place at the same time, sharing for that time the same matter, but each with its own actually or potentially distinct life-history corresponding to the abstract idea by which it is defined. On the first edition model, an individual substance which is O^1 and O^2 and O^3 and O^4 and therefore by definition an S, might also be O^5 and O^6 and therefore by definition an S^1 . ('S' and ' S^1 ' might be the same word, made ambiguous by being associated with different nominal essences.) Hence if the individual ceases to be O^{2-4} , it

ceases to be an S, but might continue to be an S¹. On the interpretation of II.xxvii.1 now being considered, however, we would be faced with a theory having the consequence that the particular S is distinct from the particular S¹. And that would mean that, if the qualities O²⁻⁴ are no longer present, the S perishes while the S¹ continues to exist.

That something is wrong with both these models should be apparent not only from their paradoxical air, but from our earlier discussions of classification, above. The name of a biological species gets its meaning, not from an arbitrary definition, but by its association with independently identifiable particulars taken to be related as members, so to speak, of the same tribe. There can therefore be no possibility of the same particular thing being a member of many species, still less a possibility of distinct particular material things of different species occupying the same place at the same time. This explanation (which, important as it is, may now be wearisome to the reader) is to be preferred to all three models for the relationship between individual and species which Locke may seem to have countenanced in the *Essay*, whether seriously or not, i.e. (i) the first edition doctrine of Epicurean particulars changing species and belonging to many species; (ii) the Boylean (or Geachian) model of existence, generation and corruption, and therefore of continuity and identity, relative to general names; and (iii) the model, suggested by other parts of the second edition discussion, of a potentially indefinite number of particulars, each individuated by a different nominal essence, occupying the same place at the same time. (How all three found expression in the same work, incidentally, might be explained as follows: the third is indeed incompatible with the first, but was developed from, and confused with, the second: while the second had been formulated by Boyle as a defence of the first, an explanation of certain apparently contrary intuitions as mere *façons de parler*.)

Yet, if we turn to the detail of the argument of II.xxvii, it soon emerges that, despite the terms of his general theory of identity, Locke's position in that chapter cannot satisfactorily be identified *either* with (ii) *or* with (iii). Nor can it be characterized unequivocally as a nominalist or conceptualist theory of identity, despite its antecedents and despite the fact that Locke himself might have wished to see it in that light.

When Boyle, with chemistry and chemical classification chiefly in mind, explained Aristotelian substantial perishing as perishing in a certain capacity, perishing as a thing of a certain kind, there is no doubt that by 'kind' he meant sort or species, so that any difference counted as specific is a difference in kind. In Locke's argument, however, the issue is considerably less clear. According to the 'general theory' of the first section, 'in this consists Identity, when the *Ideas* it is attributed to vary not at all'. The formula irresistibly suggests that the identity of a man is tied to the sortal *man*, and that different species are, within the meaning of the argument, different 'kinds'. Yet II.xxvii.2 considers only the difference between certain very general 'sorts' of substance, namely the 'Cartesian' substances, God, spirits and bodies. In the two sections following we seem to return to the suggestion that the entity we pick out, and its continuity, depends upon the specific idea, such as *horse* or *oak*. Then in the same passage we are given to understand that the idea which determines the identity of a horse or an oak is neither of these ideas, but is the generic idea of a *living body*. This last move constitutes a shift in the requirement for identity from the condition that the thing continue to satisfy a sortal

predicate, to the condition that it continue to remain alive, to embody a unifying life. (One is reminded of Hobbes' reference to 'such form as is the beginning of motion'.) It would probably be better to say that Locke confuses or conflates these two very different views, for in explaining that life is the principle of identity, he at first sticks to the formula that it should, in the case of a plant, be a life 'conformable to that sort of Plants'.

¹⁹ Given his own conception of the arbitrariness of classification, this proviso ought to make judgements of identity (or the construction of indefinitely many overlapping substantial entities with distinct life histories) an exceedingly arbitrary business. At the same time, given his dismissive explanation of all such hierarchical ordering of genus and species, it is difficult to see how he could offer any theoretical justification for picking out from the nominal essence of *horse* or *oak* the single property of being alive as alone essential to the continued existence of the individual; or for giving precedence in his account of identity to the generic idea *plant* or *living thing* over the specific idea *oak*.

It is as if Locke had merely taken over an Aristotelian respect for the logical importance of life, having rejected the ontology upon which the respect was founded. It is as if, having been forced by the standing doctrine of forms to focus attention on the significance of life and death, he brought to bear on the issue an extreme nominalist or conceptualist theory which makes identity (except identity of substance itself) relative to a nominal essence; but entirely failed to realize that nothing in that theory justifies treating loss of life as a change that is in any way special or peculiar in itself. This appearance is perhaps increased when it is recognized that the progression in his argument from the 'identity of vegetables' and the 'identity of animals' to what (for the sake of immortality and the divine judgement) he was above all interested in, 'personal identity', ²⁰ conforms to a trinity more than familiar to his seventeenth-century readers. For that progression takes in the three forms or 'Trinity of Souls' ²¹ generally attributed to a man: the generic form, involving the powers of nourishment and growth, common to all living things; the less generic form, involving sensation and self-motion, common to animals; and the specific form or difference, rationality.

Nevertheless there is implicit in Locke's argument, although not in his general theory, a philosophical justification for his procedure. For it implicitly relies upon the principle that for a property to be significant with respect to the identity of a substantial thing or object *over* time, it is necessary for it to be a property which is significant with respect to the unity and separateness of that thing at *one* time. The obverse of this principle is also on occasion assumed, i.e. that for a property to have significance with respect to diachronic identity, it is *sufficient* that it should have a certain sort of significance with respect to synchronic unity and discreteness, or distinction from other things. The interest of these principles, both of which seem to have the merit of being true, lies in their being just the sort of constraint on the formulation of criteria of identity for substances which has been postulated above. In effect the second principle was employed in the previous chapter against Locke's own ill-formed conception of a 'mass of matter'. That hybrid notion is precisely the incoherent conception of an object physically united at any moment of its existence, the continued physical or natural unity of which is taken to be only one of the conditions of its identity over time. The other, arbitrarily stipulated, condition is supposed to be its continued composition by precisely the same quantity of

matter. In contrast, when Locke turned to consider *life* he seems clearly to have recognized the link between unity and continuity. For life is taken to be responsible for, or rather to constitute, the synchronic unity of a living thing, and *therefore* to constitute its diachronic identity:

For this Organization being at any one instant in any one Collection of *Matter*, is in that particular concrete distinguished from all other, and is that individual Life, which existing constantly from that moment both forwards and backwards in the same continuity of insensibly succeeding Parts united to the living Body of the Plant, it has that Identity, which makes the same Plant, and all the parts of it, parts of the same Plant, during all the time that they exist united in that continued Organization, which is fit to convey that Common Life to all the Parts so united.²²

The close and careful relationship between this treatment of living things and the ensuing arguments about personal identity, and the extent to which the former was specifically devised as preparation for the latter, should not be overlooked. For just as life is held to be the principle of continuity of the complex animal which it unites, so, allegedly, consciousness is a principle of continuity over time just because it is a principle of unity and distinctness at any one time. One consciousness unites our ‘present Sensations and Perceptions’, and ‘makes every one to be, what he calls *self*; and thereby distinguishes himself from all other thinking things’.²³ It also unites the physical parts of the person, being the reason why ‘the Limbs of his Body is to every one a part of *himself*’.²⁴ It *therefore* serves as the principle of identity of the person over time: ‘For it is by the consciousness [which an intelligent Being] has of its present Thoughts and Actions, that it is *self* to it *self* now, and so will be the same *self* as far as the same consciousness can extend to Actions past or to come.’²⁵

If Locke had offered us a better conception of a physically united ‘body’—if, that is to say, he had not conflated the identity of a mass of matter with the identity of the matter composing it—we might have characterized his eventual position in the chapter on identity as follows: ‘The same matter or material substance may be united in just three ways at once: (i) as a material *body*, by crude physical coherence; (ii) as a living-thing of the species *man*, by life; and (iii) as a *person* (although here immaterial substance may also be involved), by consciousness. That is to say, three different principles of identity over time might somehow attach to, or be embodied in, the same quantity of matter at one and the same time (not to speak of the question of substantial continuity, the identity of the matter itself). Consequently the coherent *body* might be taken to survive the death which has destroyed the *man*, while the *person* might be held to outlast both, e.g. resurrected in a new body.’

As it is, Locke has run together the identity of the materially united, concrete body with the identity of its matter or substance. Consequently he gives no account at all of the relation between the dead body or animal and the previously living one, although there is surely *something* which exists through (all but very violent) death as one discrete thing. On a theory purged of that mistake, however, there might be just three things composed

of the same matter at the same time (given that no other uniting principles are proposed). The question that now needs to be asked is whether such a theory is any more likely to be true than the ontologically more liberal, but exceedingly implausible general theory which at least seems to be expressed in the first three sections of Locke's chapter on identity. Is it any more acceptable that just three discrete material individuals of different 'kinds' might occupy the same place at the same time, than that indefinitely many might do so?

We should first, perhaps, face up to a question of interpretation. Is the theory of identity of II.xxvii 'conceptualist' or 'nominalist' in anything like the way in which Locke's theory of classification is conceptualist or nominalist? Did Locke come to believe that the boundary-conditions of individual substances are set by our ideas or language in the same way as the boundary-conditions of species? In support of the view that he did, we can point not only to the affinity with Boyle's nominalistic explanations of 'substantial' change, but also to Locke's evident desire to present the theory of identity as a consequence of the achievement of 'clear and distinct' ideas (i.e. precise and constant definitions): 'It being one thing to be the same *Substance*, another the same *Man*, and a third the same *Person*, if *Person*, *Man* and *Substance*, are three Names standing for three different *Ideas*; for such as is the *Idea* belonging to that *Name*, such must be the *Identity*.'²⁶

Nevertheless, despite these heavy intimations of nominalism, the theory he ends up with seems realist enough. After all, the same general account of identity is supposed to cover questions about the identity of God, atoms and substance, including simple immaterial substances, all of which Locke would undoubtedly have regarded as absolute questions of ontology entirely independent of human concepts. Moreover, even compound substances, whether masses of matter, plants, animals or persons, owe their identity to principles of unity and continuity 'out there' rather than, as in the case of mixed modes, to a unity imposed by the mind. So the formula, 'such as is the *idea* belonging to that *Name*, such must be the *Identity*', need not be interpreted as a conceptualist or nominalist slogan. It could be taken as no more than a reminder that we need to know what kind of unity is possessed by the individual (whether natural or constructed) that is the subject of any question of identity.

The probability is that Locke, like many twentieth-century philosophers, was never remotely clear on this rather fundamental issue. There is undeniably a repeated suggestion that the criteria of identity for both the man and the person can simply be drawn out of our ideas of *man* and *person*, as a part of the meaning which we arbitrarily choose to ascribe to the words. In the case of 'person' there is also the insistence that the use of that 'Forensic Term' is inspired by our interest in pleasure and pain, reward and punishment. There is even an explicit analogy drawn with mixed modes in this respect.²⁷ On the other hand, the argument which finds the criterion of identity of a plant in its unitary, unifying life does not seem to take the form of a simple appeal to the idea of an *oak* 'in most people's sense',²⁸ but rather to be an appeal to mechanist theory, i.e. to the mechanist alternative to the unifying Aristotelian form or soul. It is true that the boundary and unity supplied by a single mechanical life-system, in which all the parts co-operate 'to a certain end' like the working parts of a watch, may be less absolute than the

boundary and unity purportedly supplied by a specific form. The present point is that it is nevertheless a natural boundary and unity, and that Locke's explanation of it is based on a mechanistic model of what a living thing or body *is*, rather than an analysis of the ordinary *concept* of a plant or an animal. Without that model, which was surely not the property of every seventeenth-century peasant who could identify and count plants and animals, it would be quite mysterious why *life* should be proposed as their independent principle of individuation rather than simple cohesion or, for that matter (if nominalism is thought an attractive option), some such attribute as heat or vibration. For these attributes, as much as life, can pervade for a time the whole or a part of a discrete physical object. Certainly the view of a half-dead tree which would see only the living part as constituting the biological individual is not the view of the ordinary person, as embodied in ordinary language. The un nourished, non-growing heartwood, or even an attached dead leaf, would naturally and normally be thought a part of a living tree. Locke did not consider such examples. If he had, it might have been clearer whether he saw himself as appealing to everyday ideas and language or, on the other hand, to mechanist theory as to what biological unity must be.

It is not necessary, however, to determine whether Locke's account of the identity of substances is in the end 'nominalist' (or 'conceptualist') or 'realist' in order to find an argument against the notion that three distinct material objects, distinguished by their life-histories and each unified by a distinct principle of identity, might occupy the same place, and be composed of the same matter, at the same time. The issue of personal identity is peculiar, and the argument will be restricted for the time being to the possibility of a distinction between the physically coherent, discrete body which can outlast the change from life to death, and the living-thing (e.g. man) which allegedly cannot do so. What must be said about this distinction follows from all that was proposed in Part I, above, in the argument which picked a way between Aristotle and Locke on the topic of natural classification and endeavoured to get to the root of the difference between substance and mode.

To apply a predicate like 'horse' or 'man' to an object is not to pick out something other than the independently identifiable, materially concrete, discrete thing before one, but is to classify that thing itself: to attribute to it membership of a natural class in virtue of its origin and structure. The thing's life is both a consequence or inseparable function of its origin and the continuing explanation of its structure and parts at any moment of its existence. Life is essential to the thing in so far as it is inconceivable that it (*this* thing) should have come into existence as a non-living thing. The thing exists, and exists as it does, with the structure and parts that it has, because it is alive. In that sense we can say that life is its natural principle of existence and unity. But that is not to say that when it dies the thing itself will cease to exist: merely that an explanation of the existence and structure of the thing will then refer to a life that is over. The continued unity of a plant or animal after death obviously cannot be attributed to its continuing life, but even before death the conditions of its physical coherence at any one time were laid down by previous, rather than current, life-processes. An animal or plant is not unified or given physical definition by a mysterious, instantly active life-force. There is no such force as so *constitutes* the unity of the thing that, when the force is switched off, the unity

instantly and necessarily evaporates. The sense in which life is the 'principle of unity' of a living thing is one which presupposes that 'unity' can be understood in other terms. These terms are supplied by our inescapable recognition of material coherence and discreteness. That is why the living thing can quite straightforwardly outlast its life. Life is to that extent an 'accident' of the thing. It follows that we cannot distinguish between the principle of individuation of the oak or horse and that of some essentially united mass of matter which composes it. They are therefore identical.

Life, then, cannot *constitute* the unity of an individual, but deserves to be called a 'principle of material unity' just because unity is open to explanation in each case in more primitive terms. Nevertheless, because life *is* a principle of unity, an object's being a living thing is not merely irrelevant to questions about its boundaries or identity. A stone which is enclosed by the growing trunk is not a *part* of the tree, even if it is as firmly embedded in the tree as is the heartwood. The dead heartwood has, as well as coherence, a historic causal connection with the life of the still growing part of the tree, a relation similar to the historic connection which a dead tree has with its earlier life. In normal circumstances rather less of an animal than of a tree is at any time excluded from the 'common life' (hair, hooves and so forth). These again are a part of the whole animal as an unassimilated foreign body is not. In another sort of case, what was initially a foreign body becomes a part of the individual by coming to participate in the common life, as in grafting²⁹ or transplant surgery. A plastic hip-joint, however, like a false tooth, can never become part of the individual. The relevance of life to identity also appears in certain borderline cases of identity at a time. If two (or arguably two) individuals seem not to be fully discrete, as in the case of joined twins, it is relevant (although not necessary) to our distinguishing them that each twin should have the parts necessary for a separate life, and so in a sense already lives independently of the other.

A similar consideration applies to certain trite questions of continuity. If an inanimate object, such as a block of marble, splits into two quantitatively equal parts, there is no reason to identify either of the resultant blocks with the original, and thus no reason to hold that the original individual survives as one. The same may apply to a living thing such as a herbaceous plant divided at the root. But a living thing can also be divided in such a way that only one part contains what is essential to life. The smaller part of a divided amoeba, if it contains the nucleus, may continue to live. It is then identifiable as the individual of which the larger, but dead and separated mass of protoplasm was formerly a part. The smaller part contains, so to speak, the principle of the thing's existence and survival.

These examples may go some way towards excusing the doctrine that life supplies a principle or criterion of individuation which is wholly independent of the coarse criterion of material coherence and discreteness, but they do not justify it. On the contrary, the point that the causal principle of an animal's or plant's unity has this kind of relevance to its continuing identity presupposes that animals and plants are, at least normally, coherent and discrete, i.e. that they are characteristically physically or materially unitary objects. Even in the case of animals joined at birth, a judgement that they are distinct individuals is related to a belief in their potentiality for literally separate existence. Such twins, it is true, might be thought to invite the argument that it is the sortal concept, or a 'criterion of

identity' arbitrarily associated with the sortal term, which determines the line between individuals. For it is characteristically because of our knowledge of their relationship with other members of a species that we can be sure in such a case that we are dealing with two natural individuals rather than one. Yet that is only because our knowledge of the physical nature of what is before us is normally grounded on knowledge of such a relationship. It is not because in this case (still less in all cases) the boundary round the individual member of a species is drawn by our sortal concept rather than by nature. If it does prove helpful to descend from the level of the question whether there are two *natural individuals* here or one, to that of asking whether there are two *members of the species man*, the reason is not that there is a 'criterion of identity' or counting associated with the word 'man' which is independent of material discreteness or natural boundaries. The reason is that 'man' is the general name bestowed on certain materially unitary objects to which what is before us (which may certainly belong to the species *man* even when questionably dual) is related by origin and structure. It may be helpful to know with what natural individuals it is most appropriate to compare what is before us. Attached or conjoined natural individuals may be odd ones, but by the same token (indeed, *therefore*) they will be odd *men* and *animals* too.

It is a common but mistaken thought that reference to an appropriate sortal concept can always supply, or *ought* always to supply, a definite answer, one way or the other 'with no nonsense',³⁰ to a question whether there is one individual or two. It is mistaken because such a question, when asked about substantial things, concerns a natural boundary. Consequently what our understanding of such natural boundaries cannot do, the notion of a *man, horse or plant* cannot do either. To recognize that borderline cases of identity should by no means be attributed to regrettably sloppy sortals, we need only reflect on the continuum ranging from, say, twin animals joined by a small piece of cartilage to the case of a deformed animal with eight legs, or with two heads. (Even more impressive arguments can be drawn from the world of plants.) Abnormal births can thus seem to illustrate two quite different types of borderline corresponding to the orthodox distinction referred to above between the *criterion of application* for a sortal predicate and the *criterion of identity* associated with it. We have noticed Locke's own interest in problems about the classification of monsters, problems which, in accordance with his conflation of clarity with precision, he thought could always be settled definitely by reference to a 'clear and distinct' nominal essence. He believed that when we cannot be definite about some case, it is entirely due to our failure properly to determine which criterion of application we wish to associate with the sortal term—our failure to give it a clear meaning. Such a programme of tidying up our judgements by tidying up our concepts is based on assumptions generally inappropriate to predicates having a role in natural language. But it peculiarly misconceives biological classification, which essentially depends upon the existence of underlying natural affinities, if not natural classes. It is a very similar mistake to hold that with every predicate is associated a criterion of identity which will be sharp enough, or else should be made sharp enough by 'decision', to settle every problem over how many members of the species we have, or whether we have the same individual again. For the programme of making our 'concepts' more 'precise' in this direction is likewise out of tune with the whole enterprise of

objective natural classification. Biological taxonomy in general starts from the assumption that there are *naturally* distinct individuals, to be placed as far as possible in their natural classes. With respect to either sort of natural boundaries, if they are sometimes less than perfectly definite, it will be no improvement to pretend that they can be made perfectly definite by *fiat*, by patching up 'our concepts'. It will be contrary to the essence of the whole enterprise to attempt to do any such thing. If a boundary is made sharper purely for the sake of taxonomical convenience, then it stands out as an arbitrary gloss on what is not arbitrary.

Earlier in this chapter I considered what I called constraints on the formulation of criteria of identity for substances, and accepted as such a constraint a principle implicit in a part of Locke's argument: *identity over time is tied to unity at a time*. If the rest of my argument approximates to the truth, we should need to add as a further constraint the principle that *no two particular substances can be composed of the same matter at the same time*. For it is impossible to make sense of the suggestion that two distinct principles of material unity are embodied in the same matter at once, i.e. that the same matter is materially unified in more independent ways than one. Yet this point should provoke the thought that the description of such principles as 'constraints on the formulation of criteria of identity' is profoundly misleading simply in its implication that it is our choice of criteria, even within limits, which determines the sortal's 'sense'. For the belief is thereby encouraged that the boundary to the individual substance is set, if under certain constraints, by our decision and stipulation, or by linguistic conventions. And that belief is much like Locke's false belief that we arbitrarily set the boundary to each species of substance under certain constraints (notably, under the requirement that we do not construct a criterion of application out of properties which do not in our experience coexist).

The point is that, when these 'constraints' on the stipulation of 'criteria of identity' are fully set out, they are like the famous constraint laid by Henry Ford on his customers' free choice of the colour of their cars: there is evidently no room left for stipulation or decision, individual or communal. It would therefore be better to adopt quite a different approach to the sense or meaning of sortals. Our recognition that their use is based on a presumption of natural specific boundaries (or, at least, of an underlying natural affinity or 'real essence') has led to a rejection of the doctrine that 'criteria of application' constitute part of the 'sense' or meaning of each sortal, determining its denotation. Rather, the use of the sortal, and its meaning, depends on there being a *natural class*, the boundaries of which need be neither perfectly known nor perfectly precise, of things to be denoted. We need only recognize the corollary, that the use of a sortal such as 'horse' or 'moth' equally presupposes the existence of *natural individuals* falling into a class, in order to find reason for rejecting the doctrine that another contributant to a sortal's 'sense' is something properly called a 'criterion of identity'. That is to say, any 'principle of individuation' of a particular substance is to be sought, not in a 'concept', but in reality, as the causal principle of a unity which is material and real, not imposed and ideal. It is, indeed, more than a little misleading, as well as theoretically useless, to talk of the 'concept' *horse*: almost as misleading as it is to talk of the concept *Bucephalus*.

Forms of material unity

For the sake of a name without misleading associations, let all materially unitary objects be called ‘concretions’ or ‘material concretions’. Into the class of ‘concretions’ in this technical sense fall those relatively simple and unorganized material objects, rocks, lumps of lead, chunks of salt, pieces of wood, bits of china and so forth. These I will call ‘simple concretions’.

On the Aristotelian system, simple concretions are hardly individuals at all. They are accidentally unitary and discrete, third-class individuals possessing a metaphysically lower status than things distinguished by their form, living things and artifacts. Artifacts are second-class individuals, owing their form, unity and separateness to human agency and purpose. What is perhaps surprising is that the same ordering has retained its charm long after teleology has been discarded from metaphysics. For the modern conceptualist tends to adopt the Aristotelian hierarchy as a scale of the efficiency with which different kinds of human concept perform their individuating task, a scale of the determinateness of different sorts of ‘criteria of identity’. ‘Forms’ have simply become ‘determinate concepts’.

David Wiggins’s recent account of the identity of substances (mentioned in chapter 8, above) is one that is deeply committed to the Aristotelian scale of individuality. It contains realist as well as conceptualist elements. The unity and continuity of living things are, he holds, determined by natural principles of activity, the same underlying principles as order those substances into kinds. Artifacts are ‘individuated with less logical determinacy and considerably greater arbitrariness’ than biological individuals, by principles of functioning in some way analogous to the natural principles of activity which discriminate the species of living things and their members. Below artifacts come substandard individuals designated by so-called ‘lump-mass terms’, a class which is made to include not only bars of soap and pats of butter but pools of water and pots of stew. Various as they are, these ‘lump-mass’ entities are grouped together as not following ‘the well-drilled ways of true substances’. For ‘so far as persistence and re-identification are concerned, their status is manifestly inferior to that of the fully fledged continuants that fall clearly inside the scope of a certain logical principle, called by Wiggins the ‘Thesis of the Sortal Dependency of Individuation’. In a summary formulation, this is the principle: ‘For every completely determinate continuant, there will be at least one sortal concept that it falls under and that determines a principle of persistence for it.’ ‘Chunk of metal’ and ‘piece of chalk’ pick out individuals, but are not sortal predicates. The individuals picked out by them are therefore less than completely determinate. To understand such irregulars, according to Wiggins, we simply need to understand how our purposes allow us to tolerate ‘the fuzziness of identity questions,

diachronic and synchronic', concerning them.³¹

There are three closely inter-related considerations which might be thought to reveal the indeterminateness of simple concretions. There are supposed difficulties in saying what counts as a simple concretion, supposed difficulties in counting simple concretions, and supposed difficulties in tracing them over time. In assessing these considerations it should be remembered that the existence of difficulties in certain cases will prove nothing about the status of simple concretions as a class if the same sort of difficulty arises, or can be envisaged, in cases of living things. But it would call into question the low estimation of simple concretions even more effectively if just the same very general principles as explain the identity of living things can be seen to apply to concretions in general, and to be capable of introducing a degree of order into the allegedly chaotic world of lumps and chunks.

This last suggestion might seem refuted by the mere consideration of what needs to be explained in the case of living creatures. Biological individuals quite normally survive massive replacement of matter and spectacular metamorphoses without loss of identity. It might appear self-evident that it is these achievements, denied to lumps of rock, which call for the theory that what essentially survives in living things is neither matter nor morphology, but precisely the principle or source of that normal change and normal life-history itself: its *logos* or form, so to speak. Since simple concretions have no 'normal' life-history, the argument goes, they embody no such principle of normal change. They are in equilibrium with their surroundings.³² Their survival, it is supposed, must therefore be understood quite otherwise than the survival of a living thing. Since there is nothing like the biological 'form' in question, there seems to be nothing that survival *could* consist in except the preservation of matter and/or morphology.

The argument for such a contrast goes too fast. First, the most striking achievement of living things with respect to their not being 'in equilibrium with their surroundings' is precisely their being material concretions. That the detailed course of an animal's life-history is relatively independent of specific input from outside pales into insignificance beside the animal's enduring material unity. Yet Wiggins, like many others, can offer an account of the identity of substances which never once mentions the material unity of living things. As far as his argument goes, it might be wholly accidental that animals and plants are discrete, concrete material objects. Indeed, although his theory is presented as a theory of the identity of substances, its main principles do little to differentiate substances from natural events or episodes such as thunderstorms and forest fires. The omission is self-conscious, for the difference between things and events is drawn by Wiggins on supposedly independent grounds.³³ Yet the arguments of Part I and chapters 18 and 19, above, indicate that, as we might reasonably expect, the same considerations (i.e. materiality and the kind of unity of which only material things are capable) are central both to the distinction between things and events and to the identity of individual substances over time.

The same disregard of material unity is manifested in Wiggins's list of 'lump-mass' terms, in which pools of water are ranked with bars of soap. The list appears to be constructed according to some cool linguistic principle that expressions of the form '(count-noun) of (substantial mass-term)' pick out substantial individuals, and do so in a

way to which the distinction between what is and what is not materially unitary is irrelevant. Yet reliance on such a formal ground of classification simply gets in the way of a satisfactory theory of the identity of substances. A pool of water is a 'substantial individual' in so far as water is substantial, but it is not an 'individual substance' as is a horse, or flint or piece of lead. It is unsurprising that Wiggins finds something indeterminate about the identity of such 'lump-masses' as pots of stew. Talk of a pot of stew might well be indeterminate as to whether its subject-matter is simply a quantity of stew identified through its current container, a complex of container and stuff contained, or a sloppy, extrinsically united, pretended concretion which is treated as capable of surviving replacement of its matter. In other words, their satisfaction of the formula '(count-noun) of (substantial mass-term)' is no guarantee that expressions are predicable of the same category of object, and some such expressions are often used without determinate logical form, when logical precision does not matter. Yet that does not mean that the identity of bars of soap or lumps of lead is infected with indeterminacy.

Some remarks made towards the end of the last chapter might serve to introduce a more satisfactory classification which respects the fundamental affinity between all material concretions, of whatever kind. Life, it was there suggested, owes its significance in questions of identity, not to its *constituting* the unity of a living thing, but to its being the cause or principle of the thing's *material* unity. Yet in a way life does at least help to constitute just that unity. It seems intuitively that a bullet, however firmly embedded in an animal or tree, is not a part of the living thing, whereas a graft can become so by coming to participate in the 'one common life'. But the important point is that the unity which life helps to constitute *is* material unity, not a rival to it. The animal and the bullet are not, in the sense required, *materially* united. Quite generally, the cause of material unity needs to be borne in mind in considering what material unity is in the case in question.

An animal is just one kind of material concretion, and in none is the combination of particles their merely coming to lie adjacent to one another, like marbles in a bottle. The difference between mere local conjunction and material unity is roughly illustrated by what happens when water freezes solid.³⁴ The lump of ice is unitary in that there is a special sort of relationship between its parts. If there had been a stone in the water, then, however firmly it is now wedged in the ice, it is still proper to think of it as a distinct individual rather than as a part of a materially unitary whole. Its parts are causally and historically inter-related as they are not related to the parts of the lump of ice. It is in broadly the same way that a bullet or, for that matter, a plastic hip-joint remains distinct from the animal in which it is embedded, a tree from the rock enclosed in its roots. It is unsurprising that the expressions 'lump of...', 'piece of...' and so forth are so naturally completed by single or generic stuff-sortals. Particles of metal or of stone can combine together as they do not combine with particles of wood or mud. There is a question-mark over the unity of a *lump of gold and chewing-gum* as there is not over either a *lump of gold* or a *lump of chewing-gum*. And in all such cases, whether or not living things are involved, it is facts, not 'our concepts', which cast the doubt. We cannot remove it, that is to say, by stipulatively polishing up 'our concepts'.

The analogy between the unity of living things and the unity of simple concretions can

be extended to their continuity through time. Whenever there exists a material concretion, there is necessarily a causal explanation of its origin and persistence as such. The nature of this explanation is always in turn liable to contribute to the determination of the continuity and persistence of that individual. The circularity is not vicious. Take, for example, a small lump of plasticine existing at a certain time, and a lump several times larger existing an hour later. The larger lump has been formed, let us suppose, by a child's kneading together several equally sized pieces with the original lump. The nature of this process entails that the two lumps are not the same. On the other hand, suppose that a small piece of a certain salt grows in a solution to be many times its original size. That is a proper thing to say: the same piece grows larger. Its growth is not physically very much like the growth of a living thing, but it is a characteristic natural process, a gradual accretion reflecting the nature of the salt rather than a gross combination with other pieces brought about by an external agency. It is true that lumps of salt do not as such have 'normal' life-histories or patterns of growth which are as independent of the surrounding circumstances as those of an animal. Yet this physical difference (unlike the physical difference between a pool of water and a lump of lead) does not carry with it a logical difference in their status as individuals.

These considerations should help us to deal with one source of logicians' contempt for simple concretions, namely the belief that they are peculiarly and inescapably vulnerable to a form of the paradox called *sorites*. Suppose, for example, that a small part of a piece of soap is worn away each day. At what point has the first identified piece of soap ceased to exist? If we reply, at the point at which no more than half of it remains, then we shall have the sort of trouble over the transitivity of identity that is avoidable only if we suppose that no piece of soap can survive the least loss. Yet why should we give such an answer? Why should we be uneasy about identifying a large piece of matter at one time with a small one at another time, provided that the history is appropriate? It may be assumed that in the absence of a 'form' the identity of a lump of matter can depend on nothing but the identity of the matter. But if the 'form' of the lump *is* its material unity, then we can freely agree with the common way of speaking according to which the same bar of soap continues to exist until the bitter end, the point at which no piece of soap exists at all.

There are more complicated arguments of this kind, but none which cannot be met, or at least matched by similar difficulties involving the identity of living things. We might do well, nevertheless, to probe a little deeper. There is, for example, an argument that, on pain of paradox, we *must* deny that a bar of soap which has diminished to less than half its original size is the original bar of soap. Let us suppose, for the sake of simplicity, that the wear occurs entirely at one end of the bar. At some point there is, say, a quarter of the original quantity of soap left in the bar. It is accidental (so it may be argued) that this smaller piece of soap (call it B) should have had this particular genesis. In another possible world the same piece, composed of the same particles in the same arrangement, might have been created at one fell swoop by simple division of the original bar (call the latter A). In that case, the larger of the two pieces existing after division (call this one C) would have had the greater right to be identified with A. But if A is identical with B, then there is no possible world in which A and B exist, and yet A is identical with something

(such as C) which is not identical with B. The identity of A with B cannot be contingent on the accident that some soap was gradually washed away rather than cut away in one piece. Therefore, just because of the *possibility* of a rival to B with a superior claim to be identical with A, the hypothesis that A is identical with B is unacceptable. Only where the bar of soap which remains is composed of more than half the material of the original bar (or not significantly less) is this possibility excluded. So, at least, it may be argued.

If this argument is accepted, then we would appear to be faced with a choice between three unfamiliar (and worse than unfamiliar) ontologies. Firstly, we may stipulate that, if some piece of soap exists at a certain time t and place p , then it is possible to identify a piece of soap at t and p which can survive wear just as long as it retains (significantly) over half the matter which it possessed at t . It would seem to follow that, if some piece of soap exists at t and p , then (unless the bar is brand new) there are at t and p a number of overlapping bars of soap with different life-histories; and the number of such bars will increase indefinitely with increasing wear until (close to) half the original quantity of soap is worn away. (After that things get more complicated.)

Secondly, however, we may adopt the common-sense view that no more than one piece of soap can exist at the same place at the same time. In that case, we must suppose that, when the original piece of soap becomes half-worn and so ceases to exist, it is replaced by another, and that in turn by another, and so on until no piece remains. Given steady wear, the turnover of pieces of soap will be of geometrically increasing rapidity. Another odd consequence will be that we cannot know whether we have the same bar of soap from one moment to the next unless we know just how much soap there was in the original bar when it was new.

Thirdly, we may try to avoid both these strange ontologies by stipulating that pieces of soap have the same sort of existence-conditions as Locke accords his 'masses of matter'. Any wear whatsoever means the replacement of one fleeting 'piece of soap' by another. (The reader is referred back to chapter 18 for criticisms of this unnatural expedient.) Finally, aghast, we may give up the attempt to achieve a consistent ontology, and pronounce pieces of soap 'indeterminate' individuals.

Fortunately it is possible to reject the argument which led us into these disastrous interpretations of what it is to be a lump or chunk. As we have already seen, the causal origin of any 'concretion' may be very relevant to questions concerning its identity. It is therefore a doubtful presupposition of the argument that the very same piece of soap, B, might possibly have come into existence by cutting rather than by attrition. That the surviving piece in the imagined possible world (call it B¹) is composed of the same particles as B, arranged in the same way, is evidently not by itself enough to ensure identity. Analogously, we could not bring an object back into existence by recovering its particles and replicating their arrangement.³⁵

Nevertheless a further motive for identifying B with B¹ might be at work, along the following lines. Consider Figure 2, in which P is a part of A:

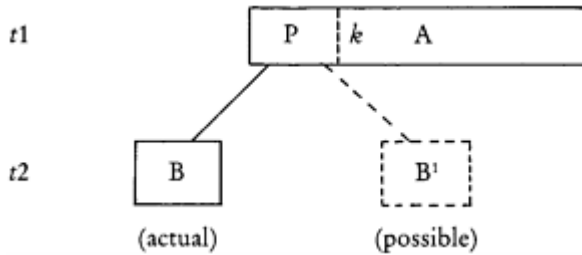


Figure 2 Wear and tear of a piece of soap

Let us suppose that B owes its present form to the gradual attrition of A (i.e. the wearing away of that part of A lying to the right of line k). The dotted lines represent the hypothesis that there might have been something, B^1 , which owed its separate existence to A's having been cleanly divided at k , rather than to attrition. Now it might be thought that B pre-existed the attrition of A as part of A, namely as P, i.e. that B is identical with P. An exactly parallel conclusion is that, if B^1 had existed, it too would have been identical with P. It would follow that actual B^1 is identical with possible B, i.e. that B might have achieved separate existence by division instead of attrition.

This argument must be fallacious. First, if it were valid, it would rule out the possibility of A's surviving any wear whatsoever, since the position of k is irrelevant to the argument. Even if P constituted nine-tenths or ninety-nine-hundredths of A, if P were identical with B, then obviously A could not be identical with B. But in any case we cannot in general identify an undetached part of something with (as we nevertheless say) that 'same' part detached. This claim is admittedly paradoxical. We might well be reluctant to agree with Aristotle that an amputated hand is not the same thing as the hand previously in use. Yet that there is a difficulty in always and in general identifying detached with corresponding undetached parts is easily shown.

Suppose a typical cat with a tail, head, legs etc. as parts. There is a very large part of such a cat, which includes every other part of it except for its tail, to which no special name is given. Let us call this part of a cat its 'untail'. An ordinary cat is larger than its 'untail' and therefore evidently not identical with it. Now suppose that our cat loses its tail (at t). Is the later cat identical with the earlier cat, or just with the earlier cat's 'untail'? Clearly the former, since a cat can survive the loss of its tail. There are three possible, now familiar ways of interpreting such a situation:

- (1) According to the view that the identity relation is always relative to a concept, we could hold that the later individual is both a cat and an untail, being both the same cat (but not the same untail) as the earlier cat and the same untail (but not the same cat) as the earlier untail.
- (2) According to the view that individuation is concept-relative, but that the identity relation is absolute, we could hold that there are two distinct individuals throughout, although at $t+1$ the cat and the untail are perfectly overlapping and

composed of the same matter at the same time. On this view, whatever survives the loss of one of its parts comes to share all its matter with (to ‘overlap’) another of its parts (that part which comprises all the remaining parts) with which it is not identical.

- (3) We could simply hold that there is one actually discrete object after t , the cat, which is not identical with the notionally discrete object before t , the untail. More generally, when an individual is divided at t , no undetached part existing before t is strictly identical with any detached part existing after t . Indeed, although the matter of an undetached part goes to compose a detached part, after t the undetached part does not exist.

The great advantage of the third option is that it avoids the curious paradoxes involved in the first, as well as the only marginally less curious paradox of two distinct material objects in the same place at the same time which is involved in the second. It may itself seem paradoxical, but is no more so in principle than a negative answer to the question, ‘does a marble sphere exist before being cut out of the block?’ The sphere is a naturally or materially discrete individual, and a notional separation cannot confer that status. What is a natural individual at one time cannot be identical with what is not a natural individual at another. If it is said that the sphere exists potentially, it should be recognized that ‘potential existence’ is not existence, but a possibility that the subject should exist.³⁶

It is not necessary to go further into the somewhat treacherous and bewildering topic of the identity or diversity of detached and undetached parts in order to see that it is unlikely to provide justification for the doctrine that simple concretions are radically less ‘determinate’ individuals than living things. Yet, having made the point that there is no such radical disparity, we need not deny that there are limits to the analogy between organic and inorganic individuals. The differences between a rock and an animal are certainly reflected to some extent in questions of continuity. In general, no doubt, an animal would not survive being rubbed away to an eighth of its size, even in the purely logical sense of ‘survive’. Moreover, the macrostructure, including the shape, of a rock is ‘accidental’ in a fairly clear sense: it does not have its origin in the ‘nature’ of the rock as the shape and complex structure of a dog has its origin in the dog’s nature. In this sense, an individual rock as such has no nature. There is therefore a sense in which an ordinary rock is not a ‘natural’ object like an animal (although perhaps a discrete crystal or mudstone is). Nevertheless there is a sense in which even a bar of soap is a ‘natural’ individual: its unity is a natural, given fact. By contrast with such secondary material objects as undetached foreheads, backs and shoulders, or with modes of any kind, it does not wait to be sliced out of reality by some human concept. It is presented to us as the unitary and discrete individual that it is. Only material things can be so presented or given; and, unless some individuals were so given, there would be no starting-point for thought of distinct individuals. All other individuals are ‘constructed’ by a kind of metaphor or analogy with natural individuals. If there were no natural individuals, there would be no constructed ones.

Here it might be objected that what is ‘given’ to us depends too much on *us* for this explanation of the primacy of substances to work. Coherence, after all, is a matter of

degree. The glass in a window-pane flows, even if slowly. The water in a raindrop coheres, even if, as it may seem to us, loosely. Even the parts of a cloud cohere to some degree. What we count as coherent depends, it may seem, on something relative to us: its resistance to human force.

To this objection it may be granted that perhaps there could be a concretion with the consistency of a cloud, although even well-defined and typical clouds do not, as it is, fit the bill. A true concretion retains its shape from moment to moment (even if it did not originally so acquire it) above all in consequence of its internal coherence. To repeat Wiggins's formula, it is not in equilibrium with its surroundings. In order to imagine a suitable example we need to postulate something like (say) a plant, existing in some improbable ecological niche, which presents as little resistance as a cloud but which has a complex structure due to its biological nature rather than to humidity, temperature, air pressure, wind speed and so forth. It is no accident that there are no such individuals, capable of acting and enduring as one thing. Yet even if there were or could be (and even if it seems an awkward question whether a drop of water is a materially unitary object) the argument of the previous paragraph is unaffected. Only substantial things can be materially coherent, and only materially coherent things enjoy an individuality which is 'given': that is to say, which is prior to their individuation by us, or relatively to our notions.

Artificial and other problematic objects

If lumps, chunks, pieces and masses of matter are, as individuals, sadly underestimated by most philosophers, artifacts, by being placed above them in the Aristotelian hierarchy, are grossly overrated. It is true that statues, spoons, invitation cards and dug-out canoes are by and large splendidly unitary objects, but they are not so *qua* artifacts. What is overrated is the role that being an artifact can play with respect to problems of unity and identity. The question is this: does its being an artifact of a particular kind, e.g. an ordinary spoon or bowl, supply a principle of unity and continuity for a substance independent of the sheer material coherence of its parts? And does their joint function unite the parts of a car or a tent in a way which is, from a logical point of view, superior to, or at any rate as good as, the way in which physical coherence unites the parts of a single piece of steel?

The question is one of fundamental importance in the theory of identity. Not only have examples like Hobbes' ship of Theseus been discussed for millennia, but any present-day conceptualist, asked to justify the distinction between a mass of matter and what it composes, would as likely as not argue in the first instance from the case of artifacts. The argument would probably be on the following lines. Suppose that a clay figure of King George III is softened and remoulded as a figure of George Washington. That means that the figure of George III has ceased to exist, while a figure of Washington has been created. Yet the same piece of clay composes the latter as composed the former. Since it has a different life history, the piece of clay must be distinct from the figure of George III (as from the figure of Washington). Therefore two distinct material objects, a piece of clay and a figure, can be composed of the same matter at the same time, or have all simultaneous parts in common.

The conclusion of this argument is paradoxical, indeed absurd. It is a paradox not only because it contradicts what we are normally prepared to say, and the principles behind what we are normally prepared to say, but because it leads to absurd dilemmas. If we are counting objects on a tray, do we count a bowl and the piece of china that composes it as two objects or as one? And if the bowl weighs a pound, are there two distinct objects which together weigh a pound, while each weighs a pound? If a bowl is chipped, are there two chips, the chip out of the bowl and the chip out of the piece of china which composes it? For that matter, if it is scratched, are there two scratches? The proposal makes nonsense of the individuation and counting of both substantial and non-substantial particulars. What then is wrong with the argument for it? That question will be answered later in the present chapter.

There is no evidence that Locke engaged in particularly sustained thought about the identity and ontological status of artifacts, and the remarks he made about them in

passing do not together constitute an unequivocal answer to the question of the opening paragraph. The analogy, in the chapter on identity, between the life of an animal or plant and a machine's 'Organization or Construction of Parts, to a certain end'³⁷ does, of course, imply an affirmative answer. It is unsurprising if philosophical mechanists were sometimes as ready as Aristotelians, if for different reasons, to compare the specific essences of living things with the functional essences of artifacts or, at any rate, machines.³⁸ Yet in the first edition of the *Essay*, in the short chapter 'Of Collective *Ideas* Of Substances', Locke had expressed the rather more interesting and suggestive thought that many artifacts, in lacking true physical unity, are conceptually, as well as literally, artificial. He proposed, in effect, that the unity of a tent is like the unity of the platoon which sleeps in it. As the artifact is in a sense a collection, so the collection is in a sense artificial:

Amongst such kind of collective *Ideas*, are to be counted most part of artificial Things, at least such of them as are made up of distinct Substances: and, in truth, if we consider all these collective *Ideas* aright, as *ARMY*, *Constellation*, *Universe*; as they are united into so many single *Ideas*, they are but the artificial Draughts of the Mind, bringing things very remote, and independent on one another, into one view, the better to contemplate, and discourse of them, united into one conception, and signified by one name.³⁹

These two suggestions are both worth exploring: first, that artifacts as such (i.e. except in so far as they happen to be truly 'distinct substances', or material concretions) are merely notional individuals, rather in the way of modes;⁴⁰ second, that their notional unity is like the unity of collections.

There are clearly reasons against classifying artifact concepts as collective concepts (although some concepts, such as *dinner service*, may be both). Tents, watches and cars have parts rather than members: a part of a part of a tent is a part of the tent, whereas a part of a member of a platoon is not a member (or part) of the platoon. Talk of the 'members' of an animal is in a different sense. The most obvious point of analogy is simply this: just as we can distinguish the real unity of a man from the notional unity of a platoon, the real unity of a tent-peg can be contrasted with the notional unity of a tent. Hence we might distinguish between car-parts and parts of car-parts, in such a sense of 'car-part' that a part of a car-part is not itself a car-part. Yet that is not enough to make car-parts, logically speaking, *members* of cars: a part of a car-part, even in this sense, *is* a part of the car of which the car-part is a part.

The analogy between artifacts and collections does, however, point to something rather deeper, which leads us back to the comparison with modes. A collection has both 'matter' and 'form'. Its form is its principle of collection, its matter is its members. When we employ a collective term we may mean to speak of the plurality of the current members of the collection, or else of that abstract or notional body which could survive even wholesale change of membership or, on the other hand, could predecease its original members. In the one sense a 'football team' can be cheerful or heavy (in the primary or literal use of these predicates): in the other sense, a team can have been more consistently

successful since 1900 than any other team in the country. The distinction is not always straightforward. A team may be described as heavier and more cheerful this year than it was last, not because its members have gained weight and thrown off their depression, but because it has new members heavier and more cheerful than the old ones. Here we may be inclined to find 'heavier' ambiguous, or some other account may be preferred. The present point, however, is in effect simply this: 'my old platoon' may refer either to a plurality of ageing men who were once fellow soldiers with the speaker, or to such an abstract object as No. 2 Platoon, C Company, 1st Loamshire Rifles, currently realized in a number of young strangers. The matter-form ambiguity of collective terms corresponds, roughly speaking, to a distinction between two sorts of entity, one material and the other notional: between a mere plurality of objects identified through their membership of the group, and the abstract group itself to which these objects belong.⁴¹

Before considering the possibility that artifact names are like collective nouns in being likewise systematically ambiguous, it may be helpful to consider some other kinds of ambiguity. One of the most familiar forms of systematic ambiguity, which supplies one of the most familiar arguments for a conceptualist approach to identity and ontology, is type-token ambiguity. The conceptualist offers to explain the type- token ambiguity of 'book' and 'dog' in terms of the possibility of our choosing either of two different 'criteria of identity' for a word given the same 'criterion of application'. In fact there is no such choice of criterion of identity to be made when a word like 'dog' is given a determinate meaning. For it is clear enough that its type-token ambiguity is logically generated in one direction: one sense is the more primitive. There exist individual animals, they are then classified into ordered taxa (named), and that hierarchical classification generates the possibility of a more abstract employment of the names of the genera. Hence 'that is a different dog from the one we saw yesterday' and even 'that is a dog' become ambiguous. A second sense is added to the primitive 'token' sense. The notion that the 'type' sense might just as well have come first is conceptualist fantasy.

In other cases of type-token ambiguity the priority of the token sense may be less clear. Since there could be no individual acts of stating without iterable content, the 'token' use of the word 'statement' automatically carries with it the possibility of 'type' use. There are, however, three features of type-token ambiguity in general which are most relevant to the present inquiry. First, it is principled, and so principled that (one supposes) it exists in some form in every natural language. Second, it is trans-categorical, involving entities of different category. Third, it is easily recognized in ordinary discourse. Because it is clearcut and overt, it is not a source of philosophical paradoxes or fallacies.

Other cases of trans-categorical ambiguity, however, can seem less principled, or less overt, or both. A plausible (if not the only possible) way of characterizing the perfectly natural speech, 'The tour of the house is nearly over. It has only to visit one more room', is to say that the speaker, no doubt unwittingly, moves from one sense of 'tour' to another: from a conception of a tour as an event, to a conception of it as a group of people. In chapter 8 above, there was mention of the somewhat indeterminate ontological status of a wave: it is conceived of sometimes as a moving quasi-substance composed of rapidly changing parts ('a body of water curling into an arched form', as one dictionary has it) and sometimes as a disturbance or ruffling of the sea. Similarly we slip easily from

‘a current of warm air’ to ‘a warm current of air’, a transition with ontological implications unlikely to be recognized by the speaker. ‘Fire’ is subject to the thing-stuff ambiguity, but also, as we have seen, to others: an individual fire may sometimes be for us a sort of compound material object such as fills a grate, with interacting parts open to rapid replacement; sometimes a particular process or activity which ‘breaks out’ and by which material things are consumed; sometimes, perhaps, it is even an event, as in ‘the Great Fire of London lasted three days’.

Another kind of case, different in that (like some type-token ambiguity) it involves straightforwardly unitary and discrete material objects, is supplied by geometrical terms. ‘Sphere’, ‘cube’, ‘triangle’ and so forth have the potentiality for a dual use as noun-predicates predicated of particulars: on the one hand as names of particular shapes realized in material things (alone or in combination); on the other hand as names of material things possessing the shape in question. We may talk of the square drawn on the blackboard or formed by a piece of string, or we may talk of a wooden square. This particular principled ambiguity, in which the same word may be used either for a substance or for a mode, supplies a simple but important analogue for the more complex and covert ambiguity which seems to be at work in at least some of our thinking about artifacts.

The suggestion that the names of artifacts are trans-categorially ambiguous, capable of naming abstract or notional particulars as well as material or substantial ones, is liable to seem shocking. If it does, then that may be partly because the two senses are not clearcut in ordinary discourse, rather as the ambiguity of collective terms is often not clearcut. It would perhaps be better to say that ordinary speech is inconsistent or confused than that the artifact-terms it employs are clearly ambiguous. The inconsistency of usage does not ordinarily matter, but it needs to be disentangled in a theory of identity. The present argument is intended to confront the inconsistency with a choice: we must either mean this or mean that. What is called for is not an arbitrary, *ad hoc* decision in order to tidy up ‘our concepts’. It is a choice between systematically and intelligibly related alternatives.

The choice we are faced with is between (i) taking the artifact name to be a term predicable of a substance or material object, in which case it is a compound or composite term with no individuating role, and artifacts are *syntheta*, or (ii) taking the name to have an individuating role, in which case it is not predicable of a substance, but of a notional or abstract object, a mode. Most philosophers presuppose that artifact names are *both* predicable of substances *and* have an individuating role. So let us consider why that is impossible.

Suppose that we define a ‘letter’ as *an inscribed material object or objects intended to convey, and capable of conveying, a message to one or more addressed individuals*. It might seem that we have built into our stipulated notion of a letter certain conditions:

- (1) that a letter is a material object,
- (2) that a letter can survive, for example, being torn up, provided that it can still convey the intended message,
- (3) that a letter will cease to exist when it becomes illegible,
- (4) that a given letter begins to exist, not at the same time as the sheet of paper which is inscribed, but when that sheet of paper is inscribed.

These conditions imply another, of course, namely the paradox (to which conceptualists are inured) that two distinct material objects, a letter and a sheet of paper, can be composed of the same matter at the same time. Yet the conditions are incompatible, since the proposed definition would place *letter* firmly in the class of composite concepts, along with *baker* and *fool*. As we have sufficiently seen, the existence-conditions of something which satisfies a composite term are not tied up with its being a member of the kind in question. If a man ceases to be a fool, nothing thereby ceases to exist but his folly. Similarly, what comes into existence when the paper is inscribed, and what goes out of existence when the ink is washed off, is an inscription: not a material object or a substance, but an accident of a substance. If it is insisted that a *letter* begins and ceases to exist under those conditions, then we can readily concede that the word 'letter' may be used for certain inscriptions, or inscribed messages. But in that usage a letter is not a material object on the same level as the sheet of paper upon which it is written. If, on the other hand, it is insisted that a letter is a material object, that can equally well be granted: but then the object we call a letter is likely to have existed before it was a letter, and might survive the obliteration of the inscription it bears. For in that sense of 'letter', the letter *is* the sheet of paper.

The same applies to the simplest kinds of artifact. Suppose that a silversmith beats a silver coin into a cup. It might be said that he has destroyed a coin and created a cup. But has he destroyed one unitary material object and brought another into being? Since all that has happened is that a piece of silver has been beaten out of one shape and into another, what has been destroyed and what has been created are neither of them material objects at all, but forms or shapes or functions of material objects. As far as the material object is concerned, it *was* a coin and *is* a cup, where the 'was' and 'is' are straightforwardly predicative. As for functions, something can acquire a function without changing at all, as a pebble off the beach can become a paperweight. Here, surely, no material object has been created.

Doubtless, in ordinary talk and thought, the dominant conception of an artifact is of a material object. Accordingly, although a bowl is a bowl in virtue of its function and shape, we allow it to lose both without holding that it has ceased to exist. We are prepared to call a crushed metal bowl just that, a bowl, although we might prefer to say that it used to be a bowl. Nevertheless there is something in ordinary ways of thinking which encourages the philosophical thought, not only that a crushed bowl is strictly speaking not a bowl, but that when a bowl is crushed, strictly speaking it ceases to exist. That will be looked at shortly, but the present point is neither to insist that letters, bowls and the like are properly regarded as material objects, nor to suggest that they are perhaps merely notional objects, but to establish that they cannot consistently be regarded *both* as material objects *and* as entities the existence and identity of which is tied to the capacity of a material object to convey a message, hold porridge or fulfil some other function. If some philosopher stipulates that 'bowl', 'letter' and so forth for him name entities of the latter sort, he can have his way as far as the present argument is concerned. But then he must take the consequences of his stipulation. He must pay for it. He is like someone who stipulates that a 'sphere' goes out of existence when a ball receives a dent, that a 'circle' begins to exist when a piece of string is suitably arranged, that a 'pattern' is destroyed

when the pieces are jumbled, or that a 'master-piece' is lost when the colours fade. There is nothing wrong with such stipulation, which in these cases leads to ways of speaking which are natural enough, indeed idiomatic. But the entities thus identified are not substances or material objects. They are abstract or notional particulars or, in Locke's terminology, 'modes'. Substances or material objects, as we have seen, have a will of their own when it comes to surviving alteration.

It might at this point be helpful to consider certain possible objections, one of which, indeed, will lead to some amendment of the account of artifact identity so far proposed. First, appeal might be made to the Aristotelian question 'What is it?' Surely, it might seem, in the case of an artifact the proper and explanatory answer to the question is simply to give the artifact name. Here, however, we should reflect on the significance of the question 'What is it?' in so far as it calls for an answer which *explains* the existence of the thing in question. It is true that a living thing owes its existence and material coherence and continuity to its being just that. It is also true, generally speaking, that if the question 'What is it?' is asked of a discrete, concrete material object which is a cup or a paperweight, then the answer, 'a cup' or 'a paperweight' will explain or cast light on the object's origins as a discrete object. It tends to be the case that an object which is a fork or spoon owes its very existence to its being a fork or spoon. That is what it was brought into existence to be. Yet that is not always or essentially the case with functional objects: in general, the object might have existed before being made into the functional object in question. When the question 'What is it?' is asked about a paperweight, the reply, 'a pebble', or 'a piece of soapstone', might be more appropriate than the perfectly true reply, 'a paperweight', if the functional characterization has in the particular case no connection with origin. Similarly the most satisfactory answer to the question when asked about a particular silver cup might be 'a silver coin beaten into the shape of a cup'. Another answer might be, 'It was a silver coin, but has been made into a cup'. Something *was* a coin and *is* a cup, where the 'was' and 'is' are straightforwardly predicative, and where there is nothing better than 'a piece of silver' for the 'something' to be. So although an artifact name very often fulfils an explanatory role in answer to the question 'What is it?', a role analogous to the role of 'horse' or 'oaktree', it is in the nature of the case not linked to the identity of the individual in the same way as those terms. Crudely, a spoon can become a shoe-horn or a pebble can become a paperweight as a goat cannot become a horse.

A simple corollary of this point answers a second possible objection closely related to the first. It might be objected to the present account that, if a concept *tin-opener* is defined as *article with the function of opening tins*, 'tin-opener' cannot be a composite term just because there is no other sortal, buried in the concept or definition, to which a criterion of identity might attach. According to this argument a term such as, 'article', 'object', 'thing' or even 'tool', if it occurs in a definition of a composite term, serves as a dummy, standing in for a more definite sortal such as 'tin-opener' itself. In the functional definition of a sortal, on the other hand, the dummy-term is entirely inert, a mere logical hook on which the expansion of the sortal is hung. As it stands, of course, this objection merely appeals to what is in question, to the principle that the identity of every individual substance is determined by a sortal under which it falls, and to the account of composite

terms as ‘restricted’ sortals. Yet, as we have seen, there is no good reason for denying that any materially unitary object is an individual substance, whether or not it is, as such, classifiable under a genuine sortal predicate. Logically, material unity comes before, and is presupposed in, classification by sortals. There is therefore no theoretical barrier to our accepting *object with the function of opening tins* as a concept of substance-plus-accident form; or ‘tin-opener’, so defined, as a composite term.

Nevertheless the objector can pursue his line of argument by tapping an important source of the notion that artifact identity is determined by function. For he can draw on examples of just the kind as led Locke to say that some artifacts are mind-dependent collections. What has so far been proposed is, in effect, that there are broadly two ways of regarding a tin-opener: either as a unitary material object which is *accidentally* used for opening tins; or as a notional object the existence and identity of which is *essentially* tied to that function. Yet a tin-opener is not necessarily a materially unified object. Tin-openers are often constructed out of a number of discrete, jointly functioning parts. In that case there *is* no material object identifiable which might ‘accidentally’ be used for opening tins. The very possibility of thinking of the collection as one material object is dependent on its capacity for opening tins, i.e. on the functional concept, *tin-opener*.

Here the difference between collective nouns and artifact names may seem significant. In its ‘material’ employment an expression like ‘the team’ refers to a plurality, and a plurality is not a single object: it has members rather than parts in the ordinary sense. This difference shows itself in matters of identity. Roughly, the team as a material plurality continues to exist as long as its members continue to exist. But a tin-opener or a car or even a tent is a material object whose existence-conditions seem both to involve more unity than that, and to leave open the possibility of loss or replacement of parts. So the way seems open for the view that the unity of such things is bestowed by the (actual or potential) joint functioning of their parts. This rival to material unity having been launched, it might seem capable of floating, as a principle of both unity and continuity, even when the artifact *is* a materially unitary object. We might then seem to have in such cases two independent principles of identity uniting the same matter at the same time, and so two distinct material objects composed of the same matter at the same time: the artifact, and the material concretion.

Such a distinction between a fork and the single piece of metal which composes it can be avoided however, if we recognize that multiplex artifacts, e.g. some tin-openers, all cars and, to date, all tents, are, as it were, *honorary* material unities, or material quasi-objects. *Because* they are composed of objects which subserve a joint function and, no less importantly, are for the most part somehow fixed together according to the intention of their maker, they are viewed *as if* they were single material things. Yet that is not to say that their functioning or capacity to function serves as an independent ‘principle of identity’. Once viewed as material individuals, they are treated as far as possible as if they had the existence-conditions of material individuals. In short, their existence is tied up with whatever second-class coherence they possess. No one but a philosopher arguing *doctrinae causa* would be inclined to say that a clock which has ceased to function has therefore ceased to exist, but it is another matter when the pieces of the clock have been pulled apart. The logical pretence is not always easy, however, and the dilemma of

Theseus' ship, discussed in chapter 18 above, illustrates some of the difficulties. A ship is repaired plank by plank until none of the originals remain. The original planks, which have been preserved, are put together to make a ship. Which of the later ships is the original ship?

On the traditional interpretation of this dilemma, the question raised is whether identity is determined by matter or by form. 'Matter' must at least be understood rather loosely here, for it is important that identifiable parts or members of the ship survive, retaining at least much of their structure and shape. If it is supposed that the replaced parts have been pulped, and that the pulp is eventually used to make new planks and a new ship, the same dilemma does not arise. That being so, the traditional interpretation suggests that the dilemma has much the same source as the following: if a football club replaces the members of its team one by one, and the original members then join together again in the team of another club, which is the original team? If that analogy were adequate, we could see the dilemma as deriving from the ambiguity of the names of artifacts. The ship which survives replacement of its members would be the notional entity realized in successively different 'matter'. The material ship, on the other hand, would be the aggregate of ship's members.

Yet such an account is more neat than satisfying, since we want the material ship to be something other than a mere plurality of parts. The force of the dilemma seems rather to spring from the indeterminacy which clings especially to the existence-conditions of a pretended, quasi-unitary material object, and which an appeal to a principle of functioning can do little to ameliorate. Can such an 'object' exist in pieces? The inclination to say that it can has more to do with the possibility of putting the pieces together again than with the likelihood that it will then work properly. Can it survive replacement of its parts? Again it is difficult to feel that involvement in the *functioning* of the whole is what makes a new part a part of the object. Some parts of artifacts have no function, or have none connected with the general function of the artifact to which they belong (e.g. the trim or clock in a car). It is how the new part fits together with the others, and the intention of the maker or repairer that it should *be* a part, which are often decisive. The role of the defining function of the artifact with respect to these questions seems minimal, and the notion that a specific 'principle of functioning' is our logical touchstone when they arise is philosophical rationalization. There is therefore no clear solution to the dilemma of Theseus' ship. The conclusion to be drawn from its consideration is perhaps no more than that artifacts are decidedly inferior individuals, with the exception of those having the good luck to be materially unitary chunks.

The logical situation of artifacts is thus somewhat messy, and our thinking in a particular case might accord with any one of (at least) the following three models.

- (1) In so far as we normally and innocently identify any artifact consisting of a materially unitary object (a single piece of metal, or whatever) with that object of which it consists, we take the artifact to exist just as long as the object exists. We ordinarily see no reason to distinguish them.
- (2) We are liable to attribute an imaginary unity to some group of objects jointly important to us because of their common purpose, their relationship in the mind of their maker, or the like, i.e. we treat them logically *as if they* constituted a

materially unitary object, a pretence which is easier in so far as they are propped, hooked, wedged, glued etc. together, but possible even if they are not.

- (3) It is possible to feel the attractions of the model according to which existence goes rigorously with function. It is easy, but mistaken, to assume that this third model differs from the second only in being logically more rigorous and principled. Such conflation in turn disguises the fact that the third model pushes artifacts out of the class of substantial objects altogether.

Misinterpretation of the second and third models can lead to the paradox of two distinct material objects composed of the same matter at the same time. The second model comes into play in such cases as the following. Suppose that a china bowl breaks and is mended, the parts being riveted together. It may seem that the single piece of china which used to compose the bowl has ceased to exist, while the bowl continues to exist; and therefore that the two must always have been distinct. Yet what is at work here, roughly speaking, is a conflict between the recognition that the unitary object which was a bowl no longer exists, and the logical pretence that the pieces of china which together serve the same purpose as the bowl constitute a genuinely unitary material object, which is a bowl. A conflict of this kind should be unworrying: there is no need to try to resolve it by a paradoxical multiplication of entities, since there is no need to take the pretence seriously.

Yet the doctrine of overlapping material objects probably gets more support from the attractions of the third model, attractions which are at their strongest in the case of representational artifacts. Hence the popular argument, set out at the beginning of the present chapter, for distinguishing a piece of clay from the statues which it successively composes. That argument should now be examined. For the purposes of criticism it can be broken down as follows:

- (1) A clay figure of George III has been refashioned as a figure of George Washington.
- (2) In this process the figure of George III has gone out of existence, while a figure of George Washington has been created.
- (3) The same piece of clay now composes the latter as composed the former.
- (4) Since they have different life-histories, that piece of clay must have been distinct from the figure of George III (as it now is from the figure of George Washington).
- (5) Therefore it has been (and is) the case that two material objects are composed of the same matter at the same time.⁴²

This argument, popular as it is (and in tune with the broadly conceptualist presuppositions of modern philosophy), is open to a number of intuitive and theoretical objections. To each objection, however, a certain theoretical reply may be given, so that the argument is best assessed dialectically. Let us suppose, then, that the initial argument gives rise to the following debate.

First objection: Taking (1) as a postulate and (3) as its implicate, (2), (4) and (5) are all false. (4) is the obvious weak point: the piece of clay palpably *was* the figure of George III, as it now *is* the figure of George Washington. We must conclude that the figure of George III survived transformation into the figure of George Washington. The two

figures are identical in the same way as (to take an example involving a phase-sortal) the boy at t is identical with the old man at $t+70$. For no one in their senses would count the figure and the piece of clay as two: numerically one object stands on the table.

First reply: In a sense the piece of clay admittedly *is* the figure of George III, and the figure *is* nothing but a piece of clay of a certain shape. But the 'is' in each case is the 'is' of composition, not the 'is' of identity or predication. It is like the 'is' of 'an animal is nothing more than organized matter'. As for counting, it is for practical reasons customary to count objects bearing the relation of composition to one another as one object, although strictly they are numerically two. The same sort of convention covers counting chips in the objects, weighing them and so forth.

Second objection: The relation between the piece of clay and the figure of George III (the relation marked by saying that the piece of clay *is* the figure) is symmetrical or reciprocal, as is identity, whereas composition is not. It cannot be said that some organized matter is composed of an animal. Nor can it be said that a piece of clay is composed of a figure.

Second reply: Whatever the truth of your last claims, as it holds between individuals, the relevant relation of composition just is the relation *is composed of the same matter as*, which is symmetrical.⁴³

Third objection: The relation *is composed of the same matter as* entails the relation of identity. That is the whole issue between us. It is a principle of part-whole logic that, if two things have all parts in common, they are identical. In its application to substances, which exist all at once and can undergo change, this principle must obviously be understood with a temporal restriction, i.e. *if A and B have all parts in common at t, then A=B*. In other words, two individual substances composed of the same matter at the same time are identical.

Third reply (i): The proposed restriction on the mereological principle is insufficient to allow its application to substances. The principle for substances should read: 'Two substances of *the same kind* which have all their parts in common at t are identical.'

Alternative third reply (ii): To apply part-whole logic to substances it is necessary to recognize that substances do *not* exist all at once. They have temporal parts. The figure of George III and the piece of clay are not identical precisely because the relation between them does not satisfy the mereological definition of identity, any more than the relation between a man and his finger satisfies that definition. The figure is a temporal part of the piece of clay. In fact, to say that they are both composed of the same matter at t is to say that the momentary phase of the piece of clay existing at t is identical with the momentary phase of the figure of George III existing at t . Things are long events, and different things can share temporal parts as easily as different events can.

Third reply (i) is, of course, nothing but a denial of what is asserted. Deadlock has been reached. The alternative reply has disadvantages which have perhaps been sufficiently discussed above. The distinction between substances and modes is obliterated, and materiality is analysed in other terms. There has been a total retreat from any attempt to understand the actual framework of our thought or the significance of the peculiarities of the category of substance. The proposed understanding of material objects as events is beset with paradox, and the present case offers a further example. For the doctrine

implies that, although the figure of George III and the piece of clay which composed it were not identical, they *would* have been identical if they had been built up and demolished together. That is to say, it implies that their identity and diversity is contingent.⁴⁴ The losses involved in such a theory can hardly be worth any supposed gain.

With respect to the whole imagined exchange, moreover, it is difficult to feel that the objector comes off the worse. His argument is considerably the stronger at the level of intuitive common sense. There is perhaps one 'intuitive' or natural judgement which he cannot accept, i.e. the judgement, which even non-philosophers might take seriously, that the figure of George III is not identical with the figure of George Washington. Yet his reaction to this judgement is not blank denial. He appeals to a plausible analogy and, in effect, to the philosophically respectable notion of a compound term or *syntheton*. The proponent of the original argument, however, has a different kind of strength. For he is arguing at the local level on behalf of a very broad philosophical view of the relationship between thought, language and reality: namely 'conceptualism' or 'idealism'. The long-standing orthodoxy of this theoretical basis is hardly a disadvantage in debate, since what is seldom questioned can seem unquestionable. Without that history, the notion of two distinct material objects' simultaneously sharing the same matter would appear a mere paradox in search of a solution, while the proposals about counting, about the 'is' of composition and about the application of mereology to substances would seem nothing but more or less specious expedients. A common, perhaps the chief philosophical purpose of the present work in its discussions of knowledge, substance and identity has been to draw from the realist tradition some equally broad theoretical backing for the kind of position adopted by my imagined objector. What, then, precisely and in detail, is wrong with the argument about the statue and the piece of clay to which he objects?

If the present account of artifacts is at all near the truth, the argument is guilty of a fallacy of equivocation. The introductory premise, (1), was expressed in ordinary language consonant with its subject's being a material object. The second premise, (2), is equivocal, in one sense being false, in the other, true. If the expression, 'the figure of George III' refers to a material object, then it is false. If it refers to that abstract or notional object, the *representation of George III*, which is embodied or realized in the material object (as the *sphere* which is a mode is realized in the *sphere* which is a substance), then it is true. Consequently, (3) is also equivocal. In one sense, the piece of clay constitutes (and is identical with) each figure in turn, and so (as the objector said) the figures are identical rather as boy and old man are identical, or as a baker who changes occupation is identical with the butcher he becomes. In the other sense, the piece of clay 'constitutes' or realizes first one figure or representation and then the other, and the representations are distinct. Turning to (4), we can see that, in the first sense, it is false that the piece of clay and the figures have different life-histories, as it is false that the boy and the man have different life-histories. Only in the second sense is (4) true. Therefore (5) does not follow. All that follows is the platitude that an abstract or notional object can be realized in a material or substantial object.

It may seem perverse to have devoted so much space to a topic not at the forefront of Locke's thinking, and on which his sayings were not only few and brief but hardly

consistent among themselves, or at best uncoordinated. There are two good reasons for having done so. First, in order to assess Locke's notion that unitary substances of different kinds (i.e. individuated by different principles of unity) can occupy the same place at the same time, it is necessary to assess the argumentation which makes such a notion acceptable to present-day philosophers. The modern arguments rely heavily and unsuspectingly on selected intuitions about artifacts in a way which was perhaps not possible when the analogy between artifacts and natural things, and so the status of artifacts as genuine objects, was more clearly recognized as problematic. The second reason for a rather careful discussion of artifacts is that it has made it possible to introduce and explain certain notions, in particular the notion of systematic trans-categorial ambiguity, which will prove essential to an assessment of Locke's theory of personal identity.

Personal identity before the *Essay*

In the first of those few entries in his journals clearly on the subject-matter of the chapter 'Of Identity and Diversity' (although the word 'identity' does not occur) Locke launched an attack on the doctrine of the natural immortality of the soul. The note begins with a statement of the 'usual physycall prooffe' of natural immortality: since matter cannot think, the soul is immaterial; since an immaterial thing is by nature indestructible (because indivisible), the soul is naturally immortal. Materialists, Locke continued, complain that animals have sensation, 'i.e., thinke', so that the same argument would prove that animals too have immortal souls. To this objection immaterialists have three possible responses: to deny (with Descartes) that animals are anything more than 'perfect machins', to allow that they do have immortal souls, or to hold that God arbitrarily annihilates their souls with their bodily deaths. Locke did not say so, but Cudworth, whose book he had just been reading, also identified these three possibilities, preferring the second as being less implausible than the first, and more economical than the third. Cudworth argued that the hypothesis of animal souls is no more disturbing theologically than the accepted principle that any substance, even matter, is naturally indestructible as such, since division is not annihilation. Locke in effect took up this point, but with a different purpose. The disputants 'perfectly mistake immortality whereby is not meant a state of bare substantiall existance and duration but a state of sensibility'. Even the 'manifestly false' doctrine that the soul thinks essentially, and 'dureing a sound quiet sleep perceives and thinkes but remembers it not', could not save the argument for natural immortality. An eternally existing soul 'with all that sense about it whereof it hath noe consciousness noe memory' simply fails to fulfil the conditions of a morally significant afterlife. To all moral effects and purposes it is dead. Since the consciousness and memory of its states are contingent activities or modifications of the soul, their occurrence at any time 'wholy depends upon the will and good pleasure of the first author': immortality is a state of grace, not the natural state of the soul.⁴⁵

Despite his use of the word 'consciousness', his emphasis on the importance of memory and other anticipations of II.xxvii, Locke's position in this early note was far from that of the second edition. It was a familiar enough point that the Christian life-to-come needs to be more than the survival of a simple immaterial substance or soul. Descartes himself had felt obliged to postulate memory after death: although one type of memory involves the corporeal imagination, the more important type, he claimed, is a function of pure intellect.⁴⁶ Henry More went further, holding that 'the immediate seat of Memory is the Soul herself': 'All Representations with their circumstances are reserved in her, not in the Spirits...nor in any part of the Body.' Indeed, 'Memory is incompetible to Matter.'⁴⁷ For More, we shall have better memories out of our bodies than we

currently have in them, for embodiment is itself analogous to the diseases known to cause amnesia. It was, then, widely agreed that, as John Tillotson (Archbishop of Canterbury and Locke's friend) put it, 'Immortality, when we ascribe it to Men, signifies two things. 1. That the Soul remains after the Body.... 2. That it lives [i.e. is active and conscious] in this separate state, and is sensible of Happiness or Misery.' It may reasonably be supposed, Tillotson added, that the souls of animals, in contrast, 'lapse into an insensible condition, and a state of inactivity'.⁴⁸

Perhaps the most original and characteristic thought of that early journal entry is the suggestion that the trite issue of immortality (i.e. whether it is natural or an arbitrary gift to naturally mortal creatures) can be settled in effect quite independently of the issue between materialism and immaterialism. Even if the soul is immaterial and naturally indestructible, immortality properly understood is due to God's special grace. Yet Locke's neutrality on the issue of materialism remained clouded by his readiness to phrase the latter part of his argument in the terms of mind-body dualism. In the following year, however, he turned to the topic of personal identity without explicit reference to immortality, and this time his starting-point seems more favourable to the materialists:

Identity of persons lies not in having the same numerically body made up of the same particles, nor if the mind consists of corporeal spirits in their being the same. But in the memory and knowledge of one's past self and actions (with the same concern one had formerly *deleted*) continued on under the consciousness of being the same person (under the certain knowledge *deleted*) whereby every man owns himself.⁴⁹

This remarkable note seems to be the first extant record of Locke's decision that consciousness of continuity is not just a necessary condition of any continuity's being ethically significant, but actually constitutes such continuity. It is as certain as these things can be that the particular context within which the thought occurred was supplied by the doctrine of the resurrection of the body, the topic of considerable controversy at the time.⁵⁰

The difficulties for a literal interpretation of the doctrine were notorious, in particular the problem set by the possibility that many particles might have been parts of more than one human body during the course of history. On one view, however, the *matter* of the body is unimportant: the same man or person would be resurrected provided only that the soul associated with the new body were the same. The ground generally advanced for this claim was that the soul can constitute the only real link, not only between a natural and a recreated body, but between the body of an infant and the body of an old man. Since the parts of a body are in continual flux, the *principium individuationis* can only lie in the immaterial soul.⁵¹ If this controversy did supply the context of Locke's note, then he was making the following proposal: even if materialism is true, it is the same consciousness rather than the same corporeal substance which unites the natural with the resurrected body. This claim was to be repeated within the much wider argument of the second edition of the *Essay*.⁵²

Soon after Locke's note of June 1683, an idea very close to his played a public role in

quite another context than the issue of immortality and resurrection, i.e. in the controversy over the doctrine of the Trinity which was later to flare up with particular heat, singeing Locke himself, in the 1690s. In 1685, John Turner, an ex-Fellow of Christ's College, published a work critical of Cudworth which set out to explain 'how it is possible for a plurality of Persons, distinct from one another, to be consistent with a Numerical Identity of Divine Substance'.⁵³

The Father, Turner claimed, is the common divine substance or nature, and a person in its own right. The Son is such a union of the divine nature with human nature as constitutes a further distinct person, while the Holy Ghost is a similar union of the divine nature with 'aetherial matter'. As a Cambridge Platonist, if a rebellious one, Turner thought of the omnipresent God as extended, having inseparable parts which are, as he put it, 'all of them acted by the same *Divine Life*, which is one self-consciousness or self-sensation running through the whole, indivisible, inseparable and *tyed* to it self, by a Unity of self-enjoyment'.⁵⁴ The thought seems to have been connected, both in content and origin, with Newton's later suggestion that space might be, in effect, the sensorium of God. At the same time, Turner argued, the compound, 'the *human nature of Christ*, vitally and personally *united* to the *Divine of God the Father*', possesses a self-consciousness which is logically distinct from the Father's. Father and Son sometimes act and think in concurrence, but at other times the Father 'acts as a Person distinctly by it self'.⁵⁵

An important element in Turner's argument was his analogy with what 'we do all of us every day experience in our selves', namely the 'vital union of an *immaterial* nature to a *material*'.⁵⁶ By this familiar union, matter 'is made to *taste* and feel it self, to become the *subject* and *seat*, either of *pleasure* or *pain*, and to concur with [immaterial substance] towards the constituting of a *common Person*, resulting from them both'.⁵⁷ There are some occasions, however, in which the immaterial substance withdraws from matter and acts on its own, namely in its purely speculative and intellectual operations. Although Turner seems to have accorded a more fundamental role to the material part of us than Descartes ever did, his account is pretty clearly indebted to Descartes' conception of the intimate, 'substantial' union of soul and body. It is, Turner claimed, even easier to conceive that two immaterial substances should achieve a similar unity-cum-duality: 'for by a *person* nothing else is meant but a *self-conscious nature*, and therefore, where there is in *two personalities* a mutual enjoyment or feeling of each *other's life*; there arises a *compound personality*'.

We know from the journal entry of June 1683 that Locke did not need to read Turner's book before arriving at something like Turner's conception of a person, but that does not establish whether he was there applying to personal continuity an idea already in the air, or whether he was doing something more original. Whichever was the case, his acquaintance William Sherlock published a work in 1690, more incendiary than Turner's, which constituted a definite link between the issue of the Trinity and the doctrines of the *Essay* itself. Sherlock advanced what is essentially Turner's explanation of the Trinity, but without the metaphysical underpinnings. For Sherlock, as for Locke, we are entirely ignorant of essences. Hence 'we know nothing of the unity of the Mind but self-consciousness...as far as consciousness reaches, so far the unity of a Spirit extends'.

Father, Son and Holy Ghost are distinguished by self-consciousness, but united by mutual consciousness. In these terms we can have an understanding of the Trinity which does not go beyond our own ideas: we need not pretend to know more about the divine essence than that God is an infinite mind.⁵⁸ It is interesting in the light of later responses to Locke's theory of personal identity that Sherlock explicitly rejected the 'Sabellian' view that the three divine persons are 'Three Modes of the same infinite God, which is little better than Three Names of One God'.⁵⁹ On the contrary, each person is substantial, 'for a Person and an intelligent Substance are reciprocal Terms':⁶⁰ a claim which, as we shall see, is open to some question.

Of the two allusions to the problems of identity over time which were present in the first edition of the *Essay*, the less interesting merely claimed that the disputes which have arisen in connection with the flux of matter, with the Neoplatonist doctrine of transmigration and with the doctrine of the resurrection of the body all go to show that, since men do not share a clear and distinct idea of identity, they cannot be supposed to share an innate one.⁶¹ In the rather longer discussion in the central sections of *Essay* II.i, however, the journal thoughts are woven into an extended criticism of the Cartesian doctrine that the soul always thinks. At the heart of this criticism is the Cartesian principle, which Locke of course accepted, that the subject of actual thought must be conscious of its thinking.⁶² If the soul were to 'have its Thinking, Enjoyment, and Concerns, its Pleasure and Pain apart' during sleep, then it would constitute a different person (or subject of consciousness) from the waking man 'consisting of Body and Soul': 'For if we take wholly away all Consciousness of our Actions and Sensations, especially of Pleasure and Pain, and the concernment that accompanies it, it will be hard to know wherein to place personal Identity.'⁶³ Locke added an illustration of this principle which is reminiscent both of Turner's multiplicity of divine persons and of his 'vital union' of soul and body: if two Cartesian body-machines were by turns united with a single Cartesian soul 'which thinks and perceives in one, what the other is never conscious of, then that soul would go to make up, with the bodies, 'two as distinct Persons...as Socrates and Plato were'.⁶⁴

Each of the passages from the journals or the first edition so far considered anticipated in some respect the argument of the second edition chapter, 'Of Identity and Diversity', yet none embodied the contention which is central to that later argument, i.e. that the clear distinction between a *person* and a *man* is made possible by the formal analogy between life and consciousness, serving as two distinct principles of unity and continuity. Yet the first edition did foreshadow the man-person distinction in a significant context (discussed in chapter 15, above) when Locke defended the thesis that the real essences 'of the Things moral Words stand for, may be perfectly known' against the objection that the subject-matter of ethical theory comprises substances as well as modes, notably the substance *man*. He distinguished '*Man* in a physical sense' from 'the *moral Man*', defined as 'a corporeal rational Being'. In the latter sense, a rational monkey would be a 'man'. Only its rationality matters, and we know in advance that its other characteristics, including its underlying constitution, are irrelevant to its status as a moral agent subject to law. In the same way a mathematician may reason about 'a Cube or Globe of Gold', although the nature and constitution of the particular body is irrelevant to the reasoning.

⁶⁵ This last analogy will prove significant, but it is now enough to see that there was already in Locke's mind, without any reference to identity, a distinction between the idea of man as a sort of animal, and the formal idea of a *rational being* which is appropriate to ethics.

These, then, were some of the threads which Locke's theory of personal identity was intended to draw together.

Locke's theory of personal identity

It is in the spirit of his earlier distinction between the natural and the moral man that Locke initiated the discussion of personal identity in the chapter 'Of Identity and Diversity'. A man, he argued, is as such like any other animal, a 'living organized Body' the principle of unity and continuity of which is its life. In common speech, 'man' simply means an animal of a certain form. It is easily shown, Locke thought, that the traditional definition, *rational animal*, falls short: an irrational animal in the 'shape and make' of a man would ordinarily be called a man, while we would count an animal which had the shape of a parrot, but which could 'discourse, reason and philosophize', as 'a very intelligent rational *Parrot*'.⁶⁶ A view to which Locke took specific objection was the doctrine (capable of either an Aristotelian or a Cartesian interpretation) that 'the *Identity* of Soul alone makes the same Man'.⁶⁷ He argued against it that we can at least make sense of the Platonic hypothesis of transmigration, according to which the same immaterial soul is successively united to the bodies of different men, or even to the bodies of men and lower animals.⁶⁸

In the light of this criticism of traditional views of what a *man* is, Locke then proceeded to offer his own careful definition of a *person*: 'a thinking, intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places'. He wished to derive the principle of individuation of a 'thinking thing' from the nature of thinking itself: it is 'that consciousness, which is inseparable from thinking, and as it seems...essential to it. It being impossible for anyone to perceive without perceiving that he does perceive'. Locke was appealing to the orthodox Cartesian notion of our necessary reflexive awareness of all our mental activity. This reflexive consciousness or apperception is what unifies 'our present Sensations and Perceptions', and 'can be extended backwards to any past Action or Thought'. Neatly, the defining attribute of a *person* supplies the principle of personal unity and, *ipso facto*, the principle of personal continuity. Even more neatly, Locke's explanation of these interconnected logical functions of consciousness came only after he had explained the supposedly similar role of *life* as the defining attribute, the principle of unity and the principle of continuity of living things.⁶⁹

The analogy between life and consciousness supplied the main framework for Locke's argument that the identity of the moral agent is conceptually independent of any particular theory about the nature, number and continuity of whatever, at the substantial level, underlies that phenomenal identity. The underlying substance may be 'Spiritual, or Material, Simple, or Compounded, it matters not'.⁷⁰ Hence neither the unity of consciousness nor moral experience and reflection on the conditions of divine justice afford any reason for belief in simple immaterial souls. On the contrary, just as life is an

organizing principle which unites a variety of 'fleeting' or ever-changing parts into one continuing animal, so consciousness is a principle which unites what is at least possibly a variety of fleeting parts into one person. To fail to see that, Locke implied, is to introduce the possibility of an alarming scepticism about personal identity, since we can rationally do no more than confess our ignorance or, at best, offer speculative hypotheses as to the substance or substances underlying thought.⁷¹ At the same time, he was not claiming that it is absolutely or metaphysically possible that the seat of consciousness is a system of fleeting parts, material or immaterial. The possibility he was concerned with was the epistemic possibility left open by our inadequate ideas: nothing in our idea of a finite self-conscious thinking thing entails either its simplicity or its immateriality. Locke was prepared to say that the most probable 'opinion' or 'hypothesis' is that the seat of consciousness is a simple immaterial soul. But it could well be, for all that we can be certain of, that a finite thinking thing must comprise material organs and 'a certain System of fleeting Animal Spirits'.⁷²

Such, then, was Locke's strategy. The precise course and details of his argument are sometimes rather less clear. Immediately after his first statement that the identity of the self is determined by consciousness, a tortuous discussion begins with an explanation of why 'it is farther enquir'd whether [the self] be the same Identical Substance'.⁷³ Was Locke thinking of inquiries actually made by his contemporaries? It is reasonable to assume that he again had in mind the controversy over the resurrection. If so, he was in effect raising the following question: if the principle of personal continuity, recognized in our ordinary concept of a person, is consciousness, why was the dispute as to the *principium individuationis* of the resurrectable self or moral agent so generally framed in terms of the identity of substance, whether soul, body or both in substantial union? An obvious answer to this question lies in the traditional principle that the identity of accidents or non-substances is dependent on the identity of substances, never the reverse. Where else is there to look for the identity of a substantial self if not in substantial identity, whether of matter or form, body or soul? Even Hobbes, in advancing a mechanistic account of animal and human identity similar to Locke's, had thought fit to distinguish 'form' from 'accident'. The defining shape of a ship is an accident, such that it (and the ship) remains numerically the same only in so far as the substance or matter is the same. But a man is defined by a 'form' or 'beginning of motion': 'that man will be always the same, whose actions and thoughts proceed all from the same beginning of motion, namely that which was in his generation'.⁷⁴ Hobbes, by borrowing the term 'form', dressed his radical departure from tradition in traditional clothes. Locke was in effect following him in claiming that life, although a mechanical process, is not a mere accident of matter but a principle of substantial unity and continuity. Yet Locke differed from Hobbes in finding another such principle in consciousness. The present passage seems to contain an admission that this latter principle commonly fails to satisfy us: we want the self to enjoy the continuity of underlying substance too.

Locke attributed such misguided concern to imperfections in the continuity of consciousness: both failures of memory and interruptions of consciousness have stimulated in us the doubt whether the same subject underlies the phenomenal train of consciousness throughout. We lose 'sight of our past selves'. In response he drew a

distinction. Doubt which surfaces as the doubt 'whether we are the same thinking thing; i.e. the same substance or no', may be reasonable, but the 'thinking thing' in this sense should not be confused with the 'thinking thing' which is the self or person: 'Different Substances, by the same consciousness (where they do partake in it) being united into one Person; as well as different Bodies, by the same Life are united into one Animal.' ⁷⁵

That the identity of substance is irrelevant to personal continuity is shown, Locke went on to suggest, by

our very Bodies, all whose Particles, whilst vitally united to this same thinking conscious self, so that we feel when they are touch'd, and are affected by, and conscious of good or harm that happens to them, are a part of our *selves*.

These words were evidently chosen with care, and they both echoed and deliberately distorted the orthodox Cartesian doctrine that a *person* is a 'vital union' of soul and body. Locke agreed that we consist in such a union of that which thinks in us with the parts of the body of which we are intimately conscious, but denied that our continuous identity requires the continuous identity of either material *or* immaterial substance. On the contrary, the role of the body as a part of the self illustrates the absence of any such requirement. If we cut off a hand, then the *substance* of the self is changed, although the self or person remains the same. ⁷⁶

Here, unusually, Locke used 'body' as we all do, as something itself capable of surviving a change of parts. But we can perhaps better sympathize with his official, 'strict and philosophical' use of 'the same body' (despite the criticisms brought against it in chapter 18, above) if we see it as being partly determined by the problem of the resurrection. Indeed, the threefold distinction between the identity of *substance* (whether the same 'body' or the same 'soul'), the identity of the *man* (or 'living organized body') and the identity of the *person* owed its existence, in all probability, to this problem and to the argument which surrounded it. ⁷⁷ Locke was able to distinguish clearly and consistently between several questions at issue: whether the resurrected body will be composed of the same particles as its natural predecessor; whether it will share the same principle of life (i.e. whether there will be the same man, in Locke's terminology); and whether it will share the same immaterial soul. His own proposal, embodied in his new (or nearly new) account of what makes 'the same person', was to set aside these questions (except in so far as a mechanistic understanding of life answers the second in the negative) in favour of the principle that the same self-conscious, rational being will be resurrected. Yet such a proposal could only be a serious *rival* to the existing proposals if it could be shown that rational consciousness is not a mere power or accident, as irrelevant to the identity of its possessor as musicality is irrelevant to the identity of the musician. It was precisely that theoretical aim which motivated Locke's insistence that life is a non-substantial principle of substantial unity, and that consciousness is relevantly just like life. The philosophical critic of Locke must decide: does the trick work?

The claim that consciousness would constitute a distinct principle of individuation whatever supposition we made about the underlying substantial basis of consciousness comes under the greatest theoretical pressure, of course, with the supposition on which

the analogy with life supervening on flux is least compelling, i.e. when what is supposed for the sake of argument is the Platonic-Cartesian hypothesis of a simple immaterial soul. How, on that special hypothesis, could a distinction possibly be drawn between the *self* which thinks and the underlying *substance* which thinks? Locke, bravely enough, faced up to the question explicitly, but the section in which he did so is singularly ambiguous.

⁷⁸ In the previous section he had again called on the alleged intelligibility of transmigration in order to drive a wedge between the continuity of the soul and the continuity of the self: only if migrating souls kept their memories would personal identity be preserved. Yet when he turned to the reverse possibility, personal continuity through a change of simple immaterial substances, his argument seems understandably to have faltered. Such a transfer of consciousness between even simple substances cannot, he claimed, be ruled out in our state of ignorance. For there is no evident theoretical reason 'why one intellectual Substance may not have represented to it, as done by itself, what it never did, and was perhaps done by some other Agent'. ⁷⁹ Yet to put the possibility this way seems an endorsement of the view that it would be a kind of illusion, and therefore a source of injustice, if such a transfer occurred:

that it never is so, will by us, till we have clearer views of the Nature of thinking Substances, be best resolv'd into the Goodness of God, who as far as the Happiness or Misery of any of his sensible Creatures is concerned in it, will not by a fatal Error of theirs transfer from one to another, that consciousness, which draws Reward or Punishment with it. ⁸⁰

The apparent incoherence in Locke's argument is only made more apparent by what immediately follows: *if* the same consciousness *can* 'be transferr'd from one thinking Substance to another', in that case 'it will be possible, that two thinking Substances may make but one Person. For the same consciousness being preserved...the personal Identity is preserv'd'. The obvious problem of interpretation lies in the question how Locke could have advanced this conclusion in the same breath as he seems to have accepted that the transfer of consciousness (i.e. of memories of what that *person* did and suffered) would involve mistake and illusion. And if the person *is* the morally significant agent, what is unjust in punishing the person for the act represented in the transferred memory? ⁸¹ Perhaps Locke couldn't help feeling that such punishment would be bad luck on the last-coming immaterial substance, which, as a subject of thought, presumably has its feelings too. A rather odd, unless ironical suggestion that there might here be a consideration counting against 'those who would place Thinking in a System of fleeting animal Spirits' is reminiscent of those advocates of the resurrection of the same body who solemnly appealed to the moral fitness of the same particles' composing the object of punishment as composed the sinner while sinning. ⁸² Perhaps, then, as one commentator has proposed, the whole argument here is ironically concessive: the transference of memory *would* involve illusion if, as some hold dogmatically, souls and selves are identical; yet it still could not be ruled out in principle by any knowledge such philosophers can pretend to have. ⁸³ On the other hand, the supposition of the wholesale transference of consciousness from one distinct simple substance to another takes us so far from the

analogy with organic life supervening on flux that it would be unsurprising if Locke had had some sceptical qualms about the rigid application of his doctrine in such a case. There seems nothing particularly ironical about his reassurance that God would not be a party to injustice.

It would perhaps be helpful at this point to review Locke's account, indeterminate and shifting as it is, of the uniting function of consciousness. At its introduction into the argument, consciousness was said to unite such items as sensations, perceptions, thoughts and actions: in general, *modes or operations* of the mind which are necessarily self-conscious. Yet in the following discussion Locke invited his reader to see consciousness as uniting *substantial parts*, and that in two ways. First, the unknown, material or immaterial, simple or complex substance which underlies and sustains a single consciousness is 'united into one Person' by that consciousness. Second, the intimate consciousness we have of the outlying parts of the body incorporates them too into 'our thinking conscious self'. There seems to be an uneasy slide here from the substance which constitutes the *subject* of consciousness to the substance which is a peculiar intentional *object* of consciousness.

Locke's example of a part of the body united to the self by consciousness is a little finger. A little finger, unlike the brain, is normally an object of feeling and may be the object of a pain, but it is not a part of the physical basis of consciousness as the brain is, or as immaterial substance was supposed to be.⁸⁴ There are accordingly, on Locke's account, two quite different ways in which consciousness unites over time. Our past thoughts and (presumably) our past bodies are commonly the intentional objects of present consciousness, whereas the past substance which thought in us (the substantial basis of consciousness) is linked to the present such substance on quite other grounds. In each case there may seem to be room for *some* analogy with life, but simply in there being two such very different grounds of unity and continuity the general analogy with life breaks down. (That is unsurprising enough, since life has no intentional objects.) Unlike life, for instance, consciousness is for ever being interrupted, and it is through the intentionality of its successive states that it is supposed to span the gaps. Indeed there are on Locke's account two fundamentally different ways in which 'one consciousness' may be either continuous or gappy: first, like one life, simply as a natural process; and, second, in virtue of intentional relationships between its elements. In the first respect, it is interrupted by deep sleep, in the second respect, by forgetting. A significant part of the difficulty in interpreting the theory is the difficulty of knowing how these two sorts of continuity are supposed to be related, in general and in particular cases.

Some see in Locke's argument quite another conception of the uniting relation involved in consciousness. The series of supposed conceptual possibilities or imaginary cases by which he seeks to break the link in his readers' minds between the self and the man, or between the self and the immaterial soul, is throughout imbued and directed by certain *moral* intuitions. Chief among these is the thought that reward and punishment simply do not make sense unless the recipient acknowledges the action in question as his own;⁸⁵ and, on the other hand, that I am 'justly accountable for any Action... appropriated to me now by this self-consciousness'.⁸⁶ Here, as others have pointed out, Locke was to some extent leaning on the connection, stronger then than now, between the

terms 'consciousness' and 'conscience'.⁸⁷ Another connection on which he put some pressure is between two senses of 'own', both of which appear in his assertion that

That with which the *consciousness* of this present thinking thing can join it self, makes the same *Person*...and so attributes to it *self*, and owns all the Actions of that thing, as its own, as far as that consciousness reaches, and no farther.⁸⁸

In its first occurrence here 'own' means *acknowledge*, while in its second it simply imports ownership in the barely logical sense. The verb 'appropriate' is another ambiguous term employed several times in the course of the chapter. Thus 'person'

is a Forensick Term appropriating Actions and their Merit.... This personality extends it *self* beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it *self* past Actions, just upon the same ground, and for the same reason, that it does the present.... And therefore whatever past Actions it cannot reconcile or appropriate to that present *self* by consciousness, it can be no more concerned in, than if they had never been done.⁸⁹

These passages have led some commentators to argue that personal continuity is not for Locke a natural relation at all: what links past to present is not cognitive consciousness of the past in the sense of *memory*, but an act of moral concern or acknowledgement which sets up a kind of non-natural, proprietary relationship. On this view Locke's notion of a person is so thoroughly 'forensic' (and so like the notion of legal property itself) that the identification of persons and questions of their continuity fall right outside the scope of ontology, being somehow concepts of pure ethics. His first edition distinction between the natural man and the moral man, discussed above, has been brought in further evidence.⁹⁰

Such an interpretation, even in so far as it is at all clear, can hardly be sustained in the face of several features of Locke's argument: first, his explicit identification of consciousness of the past with memory or recollection; second, his quasi-Cartesian explanation of 'consciousness' in terms of the necessary self-consciousness of mental operations; third, the crucial analogy between consciousness and life as natural unifying principles overlaying a flux of substance. It is memory in a straightforward sense which for Locke both ties past actions to the present and excites such moral concern as the acknowledgement of guilt. Consciousness is distinct from conscience, and prior to it. The point in associating them was to find reason for identifying the individual 'thinking thing' (i.e. that which is distinguished and united, at the experiential level, by consciousness) with the 'legal' individual, the moral agent subject to law, responsible for a past and concerned for a future. If Locke saw the class of persons as important for ethics rather than for biology, that does not mean that self-conscious rationality, which defines that class and unifies each of its members, is anything but a natural attribute.

The idea that one's body and actions are 'property' in a legal sense was not novel in

Locke's time. According to earlier Natural Law theory, each man has a sphere of what is 'his own' or '*suum*' over which he has rights. The primitive or original or natural '*suum*' includes, according to Grotius, life, body, limbs, reputation, honour and *actiones propriae*, our own actions. Thus all wrong-doing or *iniuria* (including, for example, murder) is treated as an infringement of property rights. Property in the normal sense is in turn treated as a conventional extension of the sphere of that which is one's own by nature.⁹¹ Such a model is not absurd. That actions are legal property is a natural thought, since we can sell them. To take the fruits of another's actions is to invade his property rights over his actions, to invade his personal sphere. The effect of the theory, however, is to blur the distinction between violence to the person and violence to property, not to speak of the distinction between the person and his property.

The same rhetorical assimilation of what is legally or morally 'ours' to what is physically or naturally 'ours' found expression in the *Second Treatise of Government*: not only in the famous notion that, by 'mixing his labour' with objects and land, a man extends his ownership of his own actions to material things; but also in the possibly less serious assertion that the

Fruit or Venison, which nourishes the wild Indian...must be his, and so his, i.e. a part of him, that another can no longer have any right to it, before it can do him any good for the support of his Life.⁹²

Legal incorporation is, as it were, founded on the paradigm of the physical incorporation necessary for the benefit of food to the hunter-gatherer. Yet, despite the connection that Locke's rhetoric drew between them, his view allows for quite different relations of 'ownership'. His account of the boundary-conditions of a person is an account of the most fundamental core of natural 'property' which is 'ours' in the sense that it either literally and physically composes us, or is as inseparable from us as act from agent. It is true that external property is united to us by a merely legal or moral relation, and in the most primitive cases by an act of appropriation. Our actions themselves, on the other hand, not to speak of our experiences and the parts of our bodies, are 'appropriated' to us by an entirely natural and given principle of unity, namely consciousness, rather than by some acquisitive act of acknowledgement or 'owning' on our part.

It is more than probable that there was yet another connection with previous Natural Law theory reflected in the rhetoric of II.xxvii, a connection which has to do, not with property, but with authority. Hobbes had distinguished a 'natural person', a man whose words and actions are 'considered as his own', from an 'artificial person', whose words and actions represent those of another man or institution by whose authority he acts: so to speak, whose mask or *persona* he wears. The 'author' *owns*, i.e. authorizes and takes responsibility for, the words and actions of the 'actor'.⁹³ In II.xxvii Locke was not, of course, concerned with such an external and artificial relationship as Hobbes had been, yet he employed the same term as had appeared in Hobbes' explanation of that relationship to express and characterize the internal, natural relationship between moral agents and their own actions: a relation constituted not by an act of 'owning', but by the natural unity of consciousness.

Contemporary reactions to Locke's theory

The classic philosophical objections to Locke's account of personal identity were advanced within a few years of its publication. They fall into two classes: those which are primarily formal or logical, and those which call on the possibility of forgetting. The most purely logical of all had been levelled by Robert South against Sherlock even before the second edition of the *Essay*: consciousness is the 'action' or mode of some thing (worse, it is a reflex or second-order action); but an action logically presupposes, and so cannot constitute, its agent.⁹⁴ This principle would no doubt have seemed to Locke merely to overlook the distinction between the underlying substance, whether matter or spirit, which may always be conceived to be in a state of inactivity, and a less general, compound substance such as an animal or, perhaps, a person, which is constituted by the life or consciousness which draws its parts together. The objection merely denies Locke's assertion that there are certain modifications which do supply the unity of compound individuals. Any such denial must give an acceptable account of the unity and identity of living things before it can make headway against his account of persons.

Joseph Butler commonly gets the credit for a different, if formally somewhat similar criticism, 'that consciousness of personal identity presupposes, and therefore cannot constitute, personal identity; any more than knowledge, in any other case, can constitute truth, which it presupposes'.⁹⁵ Yet Leibniz had earlier advanced much the same point, if less elegantly:

an identity which is apparent to the person concerned—one who senses himself to be the same—presupposes a real identity obtaining through each immediate temporal transition accompanied by reflection, or by the sense of *I*; because an intimate and immediate perception cannot be mistaken in the natural course of things.⁹⁶

Both Butler and Leibniz were claiming that the unity of consciousness is the consciousness of an independently unitary thing, and that our internal sense of continuity is the awareness of the continuity of a simple substance, or of something whose 'real' identity manifestly depends on the continuity of a simple substance. Yet that claim, of course, is precisely what Locke was calling into question. Once again there is confrontation rather than refutation. Both of Locke's critics deny that a mere organic continuity would constitute the 'real' identity of a substantial thing. For Butler, someone who holds that consciousness constitutes identity cannot 'mean, that the person is *really* the same, but only that he is so in a fictitious sense'. Yet our own continuity, he insists,

appears to us as a *real* identity. The extent of disagreement is therefore far too wide for Butler's famous and intuitively attractive dictum to carry much weight on its own (although we might hope to find some room for it in any final account). Certainly, after Hume and Kant, both deeply influenced by Locke, it can hardly be considered beyond all question that a unitary self is prior, whether epistemologically or ontologically, to the unity of consciousness.

Butler's argument tends to merge with those arguments which appeal to the fallibility of memory. The mere possibility of forgetting, not to speak of misremembering or illusions of memory, seems to require that an action's being mine is logically distinct from my 'consciousness' of its being mine. Illusions of memory are not discussed in the *Essay*, unless in the special context of II.xxvii.10, considered in chapter 23, above. One of Locke's defenders, Vincent Perronet, met the objection from delusions about one's past with the retort that Locke was talking about real, not illusory consciousness: a reply which could be filled out in terms of the faculty or mechanisms by which the memory-beliefs in question are acquired.⁹⁷ Locke himself offered two brief responses to an objection based on the possibility of forgetting my experiences and actions 'beyond a possibility of my retrieving them':

am I not the same Person, that did those Actions, had those Thoughts, that I was once conscious of, though I have now forgot them? To which I answer, that we must here take notice what the word *I* is applied to, which in this case is the Man only. And the same Man being presumed to be the same Person, *I* is easily here supposed to stand also for the same Person.⁹⁸

Thus the *man* can forget absolutely and remain the individual who did or suffered what is forgotten, but not the *person*. The second response relates to the Last Day. The drunkard who has forgotten 'what he did' is properly punished in this world

because the Fact is proved against him, but want of consciousness cannot be proved for him. But in the great Day, wherein the Secrets of all Hearts shall be laid open, it may be reasonable to think, no one shall be made to answer for what he knows nothing of.⁹⁹

These two responses may seem to pull in different directions. The first (like the chapter 'Of Retention', in which Locke had presented a view of memory, 'no very sure repository',¹⁰⁰ on the lines of Hobbes' account of it as 'decaying sense') suggests that men do indeed forget absolutely. When such a thing happens, however, the man becomes a different person or (as we might prefer to say) a different person becomes realized in the same man. Yet to allow *that* is to imply that either there will be many more people to be dealt with at the resurrection than there have been men, or some crimes are to go absolutely unpunished. Locke would surely have approved neither alternative. In the special case of the 'day-man' and 'night-man' he explicitly allowed for the conceptual possibility of more people than men.¹⁰¹ Yet there is no indication that he thought that

such a possibility is ever actualized, still less that it is made actual as often as we forget some of the past. What he said is therefore open to the charge of circularity, since the belief that people will be restored to the memory of all their actions presupposes a set of actions which were theirs whether or not they remember them.

The objection from forgetting was elaborated into a *reductio* by Berkeley and Reid, who each presented the possibility, in effect, that B remembers at time t^2 what A did at t^1 , and that C remembers at t^3 what B did at t^2 , but that C does not remember at t^3 what A did at t^1 . If A, B and C are *men* there is no problem: they are one and the same man. But if they are Lockean *persons*, there is an obvious contradiction. As Reid put it, C both is, and is not, the same person as A.¹⁰²

Can any sense be made of Locke's position in the face of these criticisms? It may help to make a distinction, implicit in his discussion, between what can be conceived in general and Christian faith in particular. It seems that Locke was prepared to accept, in general and in principle, that if a man (or immaterial spirit) forgets an action beyond recall, then there is no individual still in existence who is morally responsible for that action. On the basis of revelation, however, he believed that every morally significant action will be brought home to its agent at the Last Judgement, 'his Conscience accusing or excusing him'. The Lockean Christian does not pretend to know just how this will come about. Perhaps (as More and Leibniz pronounced dogmatically) experience is misleading, and every act leaves ineradicable traces in the soul (or, for that matter, uneradicated traces in the animal spirits). The point relates to Locke's systematic scepticism as to what will happen at the resurrection at the level of underlying substance. If, as his whole exposition assumes, a person is a substance in just the sense in which a horse or an oak is a substance, personal continuity must involve *some* sort of substantial continuity *somehow* analogous to the continuous flow of the substantial 'fleeting parts' of a living thing. It is not difficult, he remarked, 'to conceive, the same Person at the Resurrection, though in a Body not exactly in make or parts the same which he had there, the same consciousness going along with the soul that inhabits it'.¹⁰³ On that pretty orthodox supposition an enduring Soul is the seat of consciousness. But it is obviously less easy to conceive of an appropriate continuity on the hypothesis of materialism, i.e. a continuity which bridges the gap between death and resurrection. Locke was apparently ready to accept that, whether or not the soul is immaterial, 'he who made us at first begin to subsist here, sensible intelligent Beings...can and will restore us to the like state of Sensibility in another World'.¹⁰⁴ Yet this assurance calls for some account of how resurrection would differ from the new creation of a mere replica. It is possible that Locke was here thinking in terms of the usual conditions proposed for the resurrection of the same body (i.e. partial identity of matter and a certain conformity), insisting only that, unless the 'same consciousness' is restored in the reconstructed body, there would be no continuity of moral significance, no *personal* resurrection: 'Nothing but consciousness can unite remote Existences into the same Person, the Identity of Substance will not do it.'¹⁰⁵ So (to return to the topic of 'the secrets of all hearts') it may be that Locke would have envisaged as a possibility that the fragmented and imperfect 'persons' which, from our human point of view, we see successively realized in a forgetful (not to say, drunken) man will somehow be continued at the resurrection in an eschatologically more

satisfactory individual, embracing them all. For if the resurrection is going to involve some divine reconstruction anyway, a certain amount of constructive improvement in the direction of what is remembered might not seem out of order. However that may be, Locke's Christian believes that somehow the world will eventually turn out to be a morally more straightforward place than our everyday experience of forgetting might suggest. As a Christian philosopher who was also opposed to dogmatic metaphysics, Locke had only to present such an outcome as a conceptual possibility not demonstrably dependent on a dogmatic theory about the ultimate nature of created thinking things. If the issue which above all shaped the chapter on identity was the resurrection, the chief motive for the chapter was anti-dogmatism.

Most modern philosophers see the only possible defence or development of Locke's doctrine in the construction of a notion of psychological continuity freed from the tight link between what a person has actually done and what he can now remember. Hume's account can be read as one such development. For these philosophers, the rope of personal identity is constituted by just such overlapping strands of memory as Berkeley's *reductio* had taken as its premise. Yet such a defence fails to meet Locke's primary requirement for a satisfactory account of personal identity. It cuts his theory of personal identity away from its roots in his thought, the conception of the proper, even essential conditions of punishment and reward which he shared with many of his opponents. Difficult as it may be to maintain this connection, his most faithful apologists saw it as the last thing to be conceded, rather than the first. One such defence was that of Vincent Perronet, brought out specifically to meet Butler's objections.¹⁰⁶ Perronet insisted that the issue hangs on the forensic definition of *person*, and that the *person* who committed such and such a crime is by definition the 'proper object of punishment' for that crime:

He that has lost his Understanding, and the Remembrance of his Crimes, is yet the *same living Agent*, and may, I presume, be call'd the *same guilty Man*: But if he be not the *same Conscious Being*, or the *same proper Object of Punishment*, he is not in Mr. *Locke's* Phrase, the *same Person*.

As it stands, Perronet's defence is deeply unsatisfactory. For, with an insouciance worthy of our own ontologically insensitive times, he merely assumed that a compound term will necessarily pick out (or can be made by *fiat* to pick out) a substantial individual distinct from the individual picked out by a simple substance-term. Consequently he did not bother to mention the crucial claim that consciousness is an independent principle of substantial unity: 'And indeed' he asked, 'since [Locke's] *Idea of Person*, is a Being *actually conscious* of its own Actions, in what else can the *Identity* of such Person consist?' An obvious answer to this philosophically naïve question is that the identity of a person consists in whatever the identity of the *being* in question consists in. Perronet certainly offers no argument to the contrary.

Another of Locke's apologists, Edmund Law, was philosophically more sophisticated and adventurous than Perronet, although his starting-point was much the same. For Law, in a defence of Locke published in 1769, the word 'person' denotes

some such quality or modification in man as denominates him a moral agent, or an accountable creature.... When we apply it to any man, we do not treat of him absolutely, and in gross; but under a particular relation or precision. ¹⁰⁷

Personal identity is a matter of how far a man is the same with respect to his moral relations and obligations. The ontological implications of this thesis were made admirably explicit. Personality or accountableness ('call it...what you please') is a mode,

a creature of Society, an abstract consideration of man, necessary for the mutual benefit of him and his fellows; i.e. a mere forensic term; and to inquire after its criterion or constituent, is to inquire in what circumstances societies...have in fact agreed to inflict evil upon individuals, in order to prevent evils to the whole body from any irregular member. ¹⁰⁸

We have the concept *person*, Law thought, just because no improvement or 'instruction' ¹⁰⁹ can be achieved by inflicting evil on those who are oblivious of their faults. We properly employ the concept with respect to divine punishment because God would not pointlessly inflict evil on his creatures. General moral truths involving the idea of a person, like truths about geometrical figures, are eternal and hypothetical, without regard to actual being. But the particulars which fall under them are also in some sense 'abstract' rather than concrete. ¹¹⁰ To think that the question of the identity of the person is indistinguishable from the question of the identity of the substantial being which is rational and self-conscious is to make the same kind of mistake, Law implied, as to suppose, given a spherical object, that there is no distinction between the identity of the concrete object as a whole and the identity of the abstract sphere realized in the object. 'Person' and 'personality' are in this respect in numerous company:

specific qualities, or modes...furnish matter for a whole tribe of abstractions daily made and preserved by such terms as usually serve to denote them; whether appellatives, in order to distinguish men in their several stations and relations...to describe their character or conduct, office, etc. as parent, patriot, king, etc. or such more general, technical ones, as paternity, patriotism, kingship, etc. ¹¹¹

The argument may be less than pellucid, but Law's analogies, both with geometrical objects and with the bearers of ethically significant relations (as readers of the present volume will need no telling), are highly suggestive. He followed Locke in regarding entities of either sort as abstract particulars which instantiate eternal truths. But when Locke had compared the idea of 'the moral man' with the idea of a cube or a globe, he had refrained from drawing the conclusion that, if the latter is a mode, then so is the former. Still less did it occur to him that an expression like 'the cube' is systematically ambiguous, denoting *either* that which at some time is (or has been) cubic or, alternatively, the shape which, as a notional or abstract object, is realized in some more

substantial object or objects. Law, however, found just such ambiguity in the term 'person', and he applied his insight to the text of Locke's chapter in a way which reveals how far his interpretation had moved from Locke's intentions. We often use the term loosely, he continued, for 'the whole aggregate of the rational being'. In this 'lax, popular sense', 'person' is merely equivalent to 'man':

In which popular sense Mr. Locke manifestly takes the word, when he says it 'stands for a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking being, in different times and places.' But when the term is used more accurately and philosophically, it stands for one especial property of that thing or being.

This was certainly rough treatment for the meticulous definition from which Locke had attempted to draw out his theory of personal identity, and which had constituted part of the equally careful distinction between *person* and *man*. Yet Law evidently saw a distinction between strict and popular senses of 'person' as necessary in order to maintain the principle that the identity of the *person* is distinct from the identity of the *man who is rational*, i.e. distinct from the identity which is entirely independent of the man's still being 'what he was once before, in this single article of personality'.

Law's parallels with 'parent', 'patriot' and 'king', taken to identify bearers of moral relations as such, may remind us of Locke's view that the idea of a *father* is the idea of a relation, even though the word 'father' is predicated of one of the terms of that relation. The view may seem odd, but it is more plausible in other cases. One way of making sense of the question whether Strawson was the same *professor* as Ryle had been is to say that the identity in question is the identity of a mode, i.e. a professorial chair or (if that is any different) of a notional being, 'the Waynflete Professor of Metaphysical Philosophy', realized in this or that man or woman. What is nonsense is to suppose that the Waynflete Professor of Metaphysical Philosophy is a substantial thing of flesh and blood who is *distinct* from the man or woman of whom 'professor' is predicated. Similarly, in a case of 'multiple personality' (such as Locke's day-man and night-man) we might intelligibly ask, 'Is he the same person as he was an hour ago?', without implying that the 'person' in question is as substantial as the man. Law commended a correspondent who had reminded him that the Latin word '*persona*' commonly denotes a certain character or role which a man may possess at one time and lack at another. This unidentified writer argued that the 'difficulties that relate to a man's forgetting some actions, etc.' arise from the mistake of conflating the *persona* with the 'substance, or *homo gerens personam*'. He even claimed that the word 'personality' is a mere barbarism, since 'person' is already abstract: like 'gratitude' and 'murder', '*person...cannot be...remodified without peculiar absurdity*'.¹¹² Perhaps in all this a similar interpretation of the Trinity had had some influence. John Clendon, for example, had somewhat sycophantically seen a relevant analogy for that mystery in the fact that the Queen of Great Britain was three persons, as Queen of England, Queen of Scotland and Queen of Ireland. Like Law's correspondent, Clendon quoted Cicero: '*exiit homo personam amid, quando induit iudicis*',¹¹³

To recognize that the idea of a person is an idea of a mode gives us the remedy, Law

thought, for the ‘egregious trifling’ of Berkeley. We must, it seems, relativize personal identity to occasions, by allowing that ‘when this faculty [memory] varies, that [person] must vary also; and he become the same or not, at different times and in divers respects’. As against Butler, we can see that in ‘this case, as in some other ideal objects, to be and to be perceived, is really the same’. Similarly there is no problem over the intermittency of consciousness: ‘it would indeed be an absurdity to suppose two beginnings of existence, if the identity of a substance, being, or man were inquired into’, but in the case of a mode ‘beginning and end of existence are quite out of the question’.¹¹⁴ On this view, a person, like the state of Poland or a performance of Wagner’s *Ring*, can survive interruptions; and, like distinct geometrical figures sharing some points but not others, different persons can overlap.

The present discussion has now taken into account several broad types of theory of personal identity. Two were rejected by Locke explicitly: first, the simple view that there is no distinction between the person and the human being; and, second, the view which identifies the person with an immaterial soul or substance. (The view that identity of *material* substance is necessary for personal identity existed only with reference to the resurrection, and figured as it did in Locke’s argument only as an expository counterweight or analogue for immaterialism.) These two accounts may lie far apart or close together, depending on the associated conception of *man*. For example (to mention just the simplest conceptions available in the seventeenth century) a man may be a ‘meer piece of mechanism, a curious Frame of Clockwork’;¹¹⁵ or a combination of Aristotelian form with matter; or a vital union of a Cartesian immaterial substance with a material machine. On some versions of the last view, the identity of the whole man is directly derivative from the identity of the soul. So perhaps the most important distinction here lies between those who identified the person with the man (*without* supposing that the identity of a man is determined by the identity of an immaterial principle or substance) and, on the other hand, those who held *either* that a person is an immaterial substance, *or* that the identity of a man or person is determined by the identity of an immaterial substance. Locke’s own view, of course, was distinct from all these, since he held that a person is a substance in the sense in which a horse or oak is a substance, but that its principle of unity is consciousness rather than life mechanistically conceived. Finally, at least one of Locke’s apologists, in endeavouring to make his theory coherent while holding on to its connection with a theory of punishment, proposed that persons are not substances at all, but modes or (in some sense) abstract or notional entities.

It is now time to consider whether more recent argument has advanced the issue in any significant respect.

Neo-Lockean and anti-Lockean theories of personal identity in analytic philosophy

Whether it is a measure of Locke's continuing influence or of what is perennial in philosophy, the main theories of personal identity currently adopted by analytic philosophers cover roughly the same range as those of the early eighteenth century. Among them, the most important distinction by far holds between those who identify the person with the human animal and those who see some sort of psychological unity and continuity as grounds for separating person and animal. Of the latter, the few surviving dualists seem doomed to respond unsatisfactorily to the questions which must be answered if their theory is to have any precise content. To place spirits outside space is unintelligible, since without space individuality, or the distinction between numerical and qualitative identity, does not make sense. To locate spirits in space, on the other hand, raises insoluble questions as to how they fill it, what effects their filling a particular place has on surrounding, not to speak of coextensive, objects (i.e. what objective *difference* is made by a spirit's being *here* rather than *there*), and how such causal relations, together with purely psychological mechanisms and processes, might be supposed to fit into a general physics or account of nature. Such questions are *ex hypothesi* insoluble because even to hold that there are answers to them which make sense, but of which we are ignorant, is to imply that 'spirits' are in a broad sense material, material enough at any rate for the same distinction between the self and the body to arise all over again. It is a fashionable deduction from ontological liberalism that dualism is to be rejected on merely empirical or pragmatic grounds, i.e. just because there is now a potentially more explanatory and predictive theory in the field. It is overwhelmingly more attractive to suppose that dualism is to be rejected because it does not in the end make sense.

For whatever reason, present-day theories which distinguish between the person and the human being are not in general metaphysically dualist, but simply argue that the counting of persons proceeds on a different principle from the counting of human beings. The argument employs two main kinds of example: cases of multiple personality, in which *people* seem to outnumber biological individuals; and cases of personal survival of biological death. As an example of the first kind, we can imagine a race of two-headed giants: each head of each giant (like Locke's day-man and night-man) has its own discrete consciousness, referring to itself as 'I' and to its fellow in the second or third person. A trite example of the second kind is the easily imagined brain-transplant such that, of the two human beings involved, it is the brain-donor, rather than the brain-recipient, who will look forward to a successful operation as a kind of survival. On the other hand, if the recipient has already suffered 'brain-death', then it may be said that that

person has ceased to exist, although the human animal remains biologically alive.

The neo-Lockean takes the view that such intuitions as these, concerned as they may seem to be with peripheral and unlikely examples, nevertheless reveal the core of 'our concept' of a person. They demonstrate (it is held) that personal identity is a matter of psychological continuity, whether that is understood primarily in terms of subjective consciousness and memory (making the person what J.L.Mackie calls a 'system of co-conscious items') or in terms of more general causal and intentional links (involving character, desires, intentions, actions and so forth) such as constitute rational agency. On this basis a number of different accounts of the self have been constructed and, by some writers, conflated.

On one account, the person, as the subject of self-conscious rationality, is a physical object which is in all familiar cases co-extensive with the human being. The two objects are distinct, even when composed of the same matter, just because they are capable of having different life-histories. The general possibility of distinct yet overlapping objects is supposed to have been established by such examples as the statue distinct from the piece of clay which composes it. (The alternative interpretation supplied by Geach should be mentioned for the sake of completeness, i.e. the paradoxical thesis that A may be the same person as B and the same man as C, although B bears no identity relation to C.)

Another account in effect raises the question of why, if a unitary consciousness is what defines the boundary-conditions of a person, parts of the body which are not also parts of the seat of consciousness should nevertheless be regarded as parts of the person. As we have seen, Locke attempted an answer to this question: the body is a peculiar *object* of consciousness, and as such, it is a part of us. In support of this claim he was able to allude to the then popular notion of the vital substantial union of soul and body. The body is experienced as one with the subject of experience.¹¹⁶ So it might be argued that each person of the two-headed giant consists of those parts of the giant in which each has feeling, and which each controls. Yet that may seem an *ad hoc* justification of a mistake or confusion in the vulgar notion of a person. If personal identity is defined by continuity of consciousness, then a person (it has seemed to some) is whatever substantial object *sustains* consciousness and its continuity. Once, perhaps, it was thought that the body as a whole does so, that the hand feels and the eye sees, but now we know that the physical basis of consciousness lies in the central nervous system, and in particular in the brain. So the person (it is argued) is just a material part of the human animal, and to imagine an effective brain-transplant is to imagine a person's being transferred bodily from one animal to another. The example of the two-headed giant, it may be noticed, can be adapted to bring out the difference between the two principles by which parts of the body may be linked to a unitary consciousness: either as peculiar objects of that consciousness, or as its seat. For we can imagine that those parts of the giant's body which are the seat of one of its consciousnesses are the object of the other one.

A third neo-Lockean view is perhaps best introduced by means of two fantastic examples deployed by Derek Parfit.¹¹⁷ First, we are to suppose a machine (a 'Brain-state Transfer Device') which can replicate all the relevant features of one brain in another, imprinting the memories, character and so forth embodied in, or carried by, the first onto the second. We are also to suppose that the first is destroyed in the process. In one

version of the story, bodies are kept ready for the reception of the memories and personality of each member of a community. Every so often everyone in the community receives (as they think of it) a new body, including a brain. They regard such a 'body-change' simply as an event in their lives which is periodically necessary for good health.

In another story (of 'Teletransportation') human beings enter a machine which, before destroying them, records the total structure of their bodies. The information is transmitted to a distant receiver which constructs an exact replica from new material. The replica in each case simply takes over psychologically where the archetype left off. It is clear enough that neither the animal nor its brain has travelled. Yet we are to imagine that members of the community in which these events occur assume without question that Teletransportation is a mode of travel for the person, the referent of the personal pronoun.

The third neo-Lockean theory in the field, Parfit's own view, agrees with both these imaginary communities. It is held, roughly speaking, that what ties a person at one time to the same person at another time is not any material connection, but simply psychological (i.e. causal-cum-intentional) links. The latter are preserved through Brain-state Transfer and Teletransportation, even if by physically unorthodox means.¹¹⁸ Therefore the person is preserved. On this view, roughly speaking, a person is 'software' or 'information' rather than 'hardware': not a material object, but something realized in a material object, and capable of being transmitted from one material object to another. In other, older words, it is a notional or abstract being, a mode rather than a substance. *Person* is collected with *battle*, *dining-club* or *nation*, rather than with *horse* or *oak*.¹¹⁹

I have tried to demonstrate what might seem a remarkable instance of philosophy's relationship to its past, although the general point could be extended to many of the topics considered in the present work. Under the influence of Locke there took place in Europe in the eighteenth century what could be described as a revolution in the perception of personal identity, at any rate by a significant section of the intelligentsia. The new theory had been forged under the pressure of several distinctive concerns of the time. Chief among them, perhaps, was a concern with the possibility of a materialist Christianity, or a Christianity as undogmatic on the nature of the soul as it seemed reasonable to be on nature in general. Other, more specific problems related to the doctrines of the Resurrection, the Last Judgement and the Trinity. Few twentieth-century philosophers share these concerns, and the recent development of neo-Lockean theory is virtually independent of them (although some neo-Lockeans still see the notion of an immaterial soul as a prime target). Yet for all the transformation of our motives and, indeed, of our general philosophical theory (the various versions of idealism and reductive empiricism, for example, lie between us and Locke), the debate on personal identity has hardly moved on since the innovations of the seventeenth and eighteenth centuries. It seems that much the same temptations exist for paradox, and much the same unease is engendered when paradox is promulgated. It sometimes seems that the only significant difference lies in the replacement of the fantasies of the Trinity and the Last Day by fantasies of science fiction.

If the edge of the Lockean paradox has been blunted, however, that is unsurprising enough, given our continuing familiarity with a few philosophical texts: in particular, the arguments of Hume, Kant and Locke himself. Many of us, too, have been influenced

outside academic philosophy by those imaginative writers, from Proust to Ray Bradbury, who have taken up the thought which Samuel Clarke found so absurd and amazing, that a person is 'a fleeting transferrible Mode or Power, like the Roundness or the Mode of Motion of Circles upon the Face of a running Stream'.¹²⁰ What might seem almost more surprising from a historical point of view (given that so much twentieth-century philosophy has been directed towards bringing the category of substance and its claims to primacy into contempt) is that our ontological conscience has proved so difficult to stifle. Even the chief philosophical advocate of Teletransportation admits that he would probably be prepared to pay more for more orthodox means of travel, and even he can state emphatically that he is *not* an 'abstract entity, like a number'.¹²¹ To some, Parfit's ontological sensitivity will seem hardly acute enough. They will believe that they are not abstract entities like nations or political parties either,¹²² and that Teletransportation, as described, would be the death of the self as well as the death of the body. These ontological qualms are not mere conservatism in the face of novelty (little novelty is left after three hundred years), but are a sign that philosophy has at its core a continuing subject-matter which is not simply left behind by time.

Among the neo-Lockean theories available, the account of a person as a mode or non-substance is arguably the least incoherent. First, the theory according to which the person is a substance distinct from, but normally co-extensive with the man will have difficulty with cases of 'multiple personality'. For if it is allowed that more than one person can exist where only one man or animal exists, then it must be accepted that indefinitely many people, all of a kind, might be found in the same place at the same time. That would make the case utterly different from the argument's paradigm (the case of the statue and the piece of clay) which is generally supposed to illustrate the principle that two material objects *of different kinds* (i.e. falling under different 'sortals' or 'count-nouns') can share the same matter at the same time.

A similar difficulty exists, *mutatis mutandis*, for the theory according to which the person is substantial, being identical with the experiencing, remembering, thinking brain. A number of writers imagine cases of personal fission (or take themselves to be describing actual cases brought on by injury to the connective tissue) in which the left and right hemispheres of the brain come to sustain distinct streams of consciousness with distinct means of access to the exterior world, one side *seeing* things, for example, the other *feeling* them, one side controlling the organ of speech, the other the hand which writes. Yet there is no reason in principle why the division of 'personalities' should correspond to any such crude division of the organ of consciousness. They could simply share the same 'hardware' overall. It is possible to imagine a case of dual personality in which the scientists, in order to ameliorate the agony and inconvenience suffered by each of the 'persons' locked together in the same body, decide to transplant one hemisphere of the brain to another body. They do so, and find to their horror that each of the survivors suffers from the same dual personality. Since in that case there would be no more reason for identifying the brain, or any part of it, with one rather than another of the 'persons' involved, all would have to be regarded (as most of us regard such cases anyway) as *personae* or 'personalities' rather than distinct substantial things within the human body. In general, if it is insisted that persons are substances or material things, it cannot also be

held (at least, compatibly with existing theories of identity) that every discrete consciousness constitutes a distinct person. Why it should be supposed that any does so will then remain obscure.

The view that a person is a material object distinct from the animal with which it is normally co-extensive generates paradoxes other than the possibility that more than one material object of the same kind should occupy the same place at the same time. As has been argued at sufficient length above, even the supposition that two distinct substances of different kinds should be co-extensive or co-material is exceedingly paradoxical. When the piece of clay is scratched, then on such a view a different material object, the statue, is also scratched. We might wonder whether there is one scratch extending across two subjects or two scratches, distinct modes because their subjects are distinct. A similar line of inquiry¹²³ begins to reveal a peculiarly deep silliness in the distinction between the *person* and the *man*. Does the *man* think and talk about itself, and does it ever address others? There is no good reason for denying that it does so just as often as the person does. The uttered thought 'I shall survive' is therefore thought and uttered by the man and the person at once. If man and person are distinct, then one psychological or linguistic event manages to bear a dual propositional content with dual self-reference.¹²⁴ Or are there perhaps two utterances, two events having just the same causal significance but nevertheless distinct? However that may be, if this utterance (or these utterances) occurred before a brain-transplant operation, then it might be that the thought and speech of the *person* were true, while those of the *man* were false. The surgeon who economically reassured both beings at once, 'You will survive', would be deceiving the *man*, giving reasonable expectation to the *person*. But then we might wonder who or what 'the surgeon' was. Should we perhaps speculate that the *man* is lying to its counterpart, and the *person* being frank to its fellow-person? Nothing, of course, would determine the truth or falsity of such a claim. Locke argued, in discussing the broadly related issue of forgetting, that 'we must here take notice of what the Word I is applied to'. He arbitrarily decided that, when talk is of my forgetting absolutely, 'I' refers to 'the man only'.¹²⁵ But that is to suppose that the 'man' could be thinking and talking about itself while the 'person' is keeping silent; as if the piece of clay could be chipped without its co-material statue being chipped.

What is the object of self-reference if not the self or person? Something may beguile us into distinguishing the self from the man, but the logical consequence of that division is to divide both self-reference and the object of self-reference, a deeply absurd and unsatisfactory move. Here it is that the rival doctrine that the identity-relation itself is relative might seem to receive a piece of support, for it might well seem better to avoid such bifurcations by taking our imaginary pre-operational utterance, 'I shall survive, although my body won't', as expressing the single thought of a single being, equivalent to 'I shall survive *as a person* or subject of consciousness, but not *as an animal* or human being'. One such advantage, however, does not justify a theory, and it will shortly be seen how this point can be accommodated within a less paradoxical approach than Geach's.

The kind of deep absurdity involved in the thesis of distinct co-material substances is not altogether avoided by a theory which distinguishes the person from the animal but,

explicitly or in effect, judges the former a 'fleeting, transferable mode' rather than a substantial object. ¹²⁶ Yet first we should take note of the question whether a psychologically unitary *persona* is the kind of thing which can be a subject of thinking or speaking at all. To suppose that it is so is like supposing that my consciousness perceives that *p*, or that my personality expressed the fear that *q*. Let us assume, however, that a 'person' is a non-substance with respect to which such a way of talking is acceptable. That would make a person analogous to, say, the Monarch of England conceived of as a ruler who has had legal existence, with brief intervals, for over a thousand years and who can, in certain conditions, properly be said to arrive at a decision or to make a pronouncement. These conditions would have to consist in certain circumstances surrounding appropriate utterance on the part of someone less notional: currently, Elizabeth Windsor. Similarly, in a case of multiple personality, whether it is 'Mary' or 'John' who is now thinking and speaking will depend on which previous thoughts and utterances are 'psychologically connected' with the present thoughts and utterances of a 'man'. It is the 'man' who will be thinking and speaking in the primary sense. No argument could make it plausible that only 'persons' in this sense think and speak and act, i.e. that human animals never do these things. It therefore seems that this way of dealing with persons will again be beset with problems of self-reference. Whenever a person thinks or speaks or acts, a man thinks or speaks or acts. Moreover, whenever a person has a conception of itself, a man has just the same conception of itself. When a person has (on this view) a true conception of itself as a 'fleeting, transferable mode', a man has a false conception of itself as something that is not a material object. Whenever a man truly thinks of itself as a substantial thing of flesh and blood, then a person falsely supposes that it is quite a different sort of object from 'the French Socialist Party'. ¹²⁷ Parfit, in offering us this last analogy, sees himself as offering a *better* way of thinking about ourselves than our ordinary way. Since he has rational animals to convince as well as 'persons', it seems that the improvement can be achieved only at the cost of inculcating false-hood; unless perhaps he can persuade us all, men and persons, to avoid self-referential thoughts altogether.

In the face of neo-Lockean theory about how we ought to think of ourselves, it is worth considering how in ordinary life, away from the study and lecture room or, for that matter, the pulpit and the philosophical novel, we do *in fact* think of ourselves. And one might begin by asking how we experience ourselves, prior to philosophical theory. ¹²⁸ In Volume I, Part III, it was argued that in order to do justice to the intentional content of sensation it is necessary to recognize that sense-experience does not consist in five (or any number of) disparate streams of information, which have to be associated by custom or inference or speculation or, for that matter, by innate disposition. The 'fields' or *contents* of different senses are themselves integrated in respect of the presentation of space. We see and hear things, for example, from the head, or even particular parts of the head, which are not themselves seen or heard. Hence the possibility (*pace* Berkeley) of seeing *everything* upside-down. The sense of touch is intimately tied in with our awareness of our own bodies, so that, unless in some strange state of disorientation, we could hardly learn anything about an external object without at the same time learning about the relation of that object to ourselves. (Disorientation itself has spatial content.)

That we should actually hear the direction of sounds or see things a particular way up (reversible by lenses) becomes unintelligible if it is assumed that sensory content cannot span the distinction between senses, that their fields cannot be integrated at the level of sensation itself. Yet the possibility of such intersensory perception consists in the fact that our bodily sensations of ourselves as objects extended in space (a class of sensations which itself may derive from a variety of sensory organs and processes, pro-prioceptive, vestibular and, perhaps, others) are integrated with the deliverances of each of the senses either quite straightforwardly (as taste is integrated with sensations of the mouth, touch with awareness of one's limbs) or slightly less directly (as we see things from, and in their relation to, the upper part of the face).

In other words our experience of *ourselves* as being a material object among others essentially permeates our sensory experience of things in general. To try to imagine experience in general free from self-awareness is like trying to imagine seeing something, but not from any point of view, or tasting something or feeling warmth, but in or at no part of the body. It was suggested in Volume I that this sensory self-awareness is bound up both with our capacity for physical action and with the having of sensory images. Perhaps all three underlie our capacity for 'intellectual' thought, such as the reasoning which may be advanced by our uttering, or imagining ourselves uttering mathematical formulae. Such proposals constitute, in effect, an argument for two linked, characteristically Kantian thoughts, although without the idealist conclusion which Kant drew from them: first, that the unity of apperception or consciousness essentially involves a unity of content; and, second, that the physical or material self is essentially involved in that double-sided unity. It is involved neither simply as the postulated subject of experience, nor simply as an ever-present object of experience, but as the *presented* subject of experience and action. The experience of pain, for example, since pain is necessarily in some part of the body, presents the bodily self from within as the subject of pain as surely as the sensation of colour presents some object or area from without as the subject of colour. Something similar is true of every bodily sensation; and, since all other forms of sensory awareness are necessarily integrated with awareness of the body, it is true of all sensory experience whatsoever.

Nevertheless, centuries of philosophical doctrine, including Kant's own, can impede our taking this kind of self-awareness seriously into account in a theory of personal identity. Somehow even anti-dualist argument has taken off from the presuppositions of dogmatic dualism. For Descartes, in perceiving anything external I have at the same time knowledge of myself as the subject of the reflexively known act of perception. Yet this apprehended self is not the self which enters into the content of sense-experience, the material self, but a self whose essence would be veridically perceived by the pure intellect even if none of our sense-experience were veridical, and even if there were no objects in space at all. Those who reject this claim that we are reflexively aware of an immaterial, unitary self have nevertheless tended to accept the way in which the dualist has set up the question of knowledge of the self. They ask (as Locke and Hume asked) whether reflexive or second-order consciousness of first-order mental activity carries with it knowledge of the nature of the thing which thinks, the subject of the first-order mental activity. The content of the first-order activity has thus seemed more or less

irrelevant to the question of the nature of the self. It is true (as we have seen) that the body was treated by Locke as the object of a peculiarly intimate kind of first-order experience which somehow weds it to the subject of mental states; yet in this too he simply followed Descartes' notion of the substantial union of soul and body. For both philosophers the point was in effect a mere rider to the theory of the core-self, and hardly more than an explanation of the merely popular notion of the self.

A closely related distinction between a core-self and a bodily self emerges in Kant's critique of dualism. In an argument deeply indebted to Locke, he opposed the view that an immaterial thinking substance is presented in experience as the enduring subject of experience. What is automatically or tautologically given in continuous experience is the continuity of a purely formal subject, the 'I' of apperception or self-consciousness; but that takes us no further than the proposition that the experience is continuous. For all we know, it might be that the continuous consciousness is being passed on from one thinking substance to another.¹²⁹

At the same time Kant recognized that a 'thinking being', the *man* who is an object of the outer senses, is experienced as an enduring object throughout our life. As Kant remarked, 'this is very far from satisfying the rational psychologist who undertakes to prove from mere concepts the absolute permanence of the self beyond this life'.¹³⁰ Yet Kant seems to have been no less inclined than the dogmatic dualist to underrate the claims of the bodily or empirical self to be, and to be presented as, the subject of experience and consciousness generally. His philosophical interest centred on the transcendental or pure subject of the states known through 'inner sense', a thing in itself whose nature can be approached only indirectly and imperfectly through our apprehension of the moral law. If the transcendental subject and the bodily self are ultimately to be brought together, that, it seems, is only because Kant envisaged the possibility that the bodily self is the outward appearance of something not in itself material.¹³¹ Only such a 'hypothesis', he claimed, can justify 'the common expression that men think'!¹³²

This traditional emphasis on the pure subject of thought as something at least conceptually distinct from the bodily self plays an almost unargued role in the current popularity of neo-Lockean theories. Thus Shoemaker can claim without qualification that when 'one is introspectively aware of one's thoughts, feelings, beliefs and desires, one is not presented to oneself as a flesh and blood person, and one does not seem to be presented to oneself as an *object* at all'. Again, in remembering one's past actions and experiences, 'one's past self, the subject of those actions and experiences, does not enter into the content of one's memory in the way other persons do'. The grounds given for these bold, implausible judgements seem to be as follows. If the subject entered into the content of experience and memory, then I could in principle misidentify that subject as someone other than myself: yet, when I am having an experience, it is impossible that I should know that *someone* is having it and fail to know that that person is myself.¹³³

Shoemaker's reasoning is evidently fallacious. If it were valid, we could as well argue that my *body* is not presented in a peculiar way in sense-experience, as no other body is presented, i.e. it could as well be argued that there are no bodily sensations. For it could be argued that, if my body were so presented, then I could be aware of the disposition of

a body and yet be mistaken as to which body is the object of my awareness, which is absurd. In fact, of course, the impossibility of misidentifying which body I am peculiarly aware of gives no grounds for denying that I have a peculiar sensory awareness of my own body. We do have bodily sensations such as (most spectacularly) pain, and we cannot intelligibly wonder whether the body so presented is our own or someone else's. It is, indeed, the physical disposition of *myself* of which I am aware in having bodily sensations, and which, as we have seen, enters in some way into the content of all experience. As much as the dogmatic dualist, the neo-Lockean characteristically ignores or sets aside the bodily self. Of course, as Hume and Kant found, no *other* object is presented as the subject of experience, and so the way is clear for the neo-Lockean to construct an alternative 'self' by means of a relation between experiences. Yet we do not need this second self: the first, given, substantial, material self is all that is required. Of course, we can improve our conception or understanding or knowledge of the substantial self, just as we can improve our conception or knowledge of any experienced material object. Yet that cannot involve our adopting a principle or 'criterion' of identity for it which abstracts from its materiality.

Some of our earlier conclusions can be brought to bear on the question of why 'consciousness' or psychological unity cannot be the principle of unity of a substance. Locke, as we have seen, presented it as such through the medium of the analogy with life, having satisfied himself that life is a principle of unity and continuity which is quite distinct from mere coherence. It was argued in chapters 18–20, above, that the premise is false. The matter of a horse or tree is not collected in two distinct ways, by crude material coherence on the one hand, and by life on the other. On the contrary, we can think of life as a principle of unity of individual substances just because individual animals and plants are in general materially unitary things which owe their existence as such to their living or having lived. Even in those special cases (such as twins joined at birth) in which life may seem to serve as a principle of individuation independent of material unity, the appearance is an illusion. There is no principle of counting lives which is independent of counting materially coherent and discrete objects which are alive.

As for the supposed analogy between the life of the body and the life of the mind, it will readily be seen that consciousness has no such causal connections as has life with the existence, material unity and continuity of the animal which is the subject of consciousness. That individual was brought into existence *qua* living thing, but not *qua* conscious thing. Moreover, in virtue of that 'unity of consciousness' which consists in intentional relations between psychological states, we can operate with a principle of counting 'consciousnesses' which is independent of the counting of materially unitary objects. (That is so even though an apprehension of its bearer as a materially discrete thing is a deep feature of all such consciousnesses of which we have knowledge and may, despite fantasy, be an essential condition of psychological unity *tout court*.) Thus our two-headed giants, each a typical member of its kind, are not on a par with conjoined twins. Unitary consciousnesses are neither individual substances themselves nor the grounds of individuation of the substances which bear them. To insist that we can simply stipulate that psychological unity serve as the principle of individuation of a sort of substances, *persons*, is as wrong-headed and ineffectual as it would be to insist that we

can arbitrarily make its heat the principle of individuation of something as substantial as a hot-cross-bun. If we do make the latter stipulation, we shall find that the individual we have picked out is no more substantial than the state of affairs describable as the hot-cross-bun's being hot.

Here, as it has been suggested more than once above, Locke's own analogy with a cube illustrates the constraints on our power of stipulation: a 'cube' may either be a cubic object with identity-conditions independent of any choice or thought or concept of ours; or it may be something which exists just as long as an object is cubic. In the first sense of 'cube', a cube is a substantial object: in the second, it is a mode or notional being the individuality of which depends on abstraction. We cannot *both* stipulate identity-conditions appropriate to the second sense *and* stipulate (or expect) that such a cube is a substance. The ambiguity of 'cube' is hardly news, but we have seen how a similar, less overt ambiguity can cause trouble in the case of representational objects, such as pictures and statues. As our case of 'multiple personality' illustrates, if we choose to count 'persons' solely with reference to unity of consciousness they degenerate into *personae*: 'John is speaking' becomes a way of saying that Richard Roe is speaking in the role of 'John'. One misleading feature of such a case is that it might be an immensely convenient *façon de parler* to enter into the play, so to speak, and to talk *as if* 'John' really were a substantial object distinct from 'Mary'. (It might be similarly convenient, in the event, to pretend that real people can travel by Teletransportation.) But ontology, as I have tried to show throughout the present work, is not just a matter of convenience.

The crucial thought-experiments, however, might seem to be those involving personal survival. It is difficult not to feel that, in a successful operation to transplant a functioning brain from one body to another, it would be the brain-donor who would survive in the egoistically important sense. In that case, we might seem forced by the logic of identity to concede that the person who survives is distinct from the man or animal who does not, and that personal continuity is constituted by psychological continuity.

The paradoxicality of this last conclusion, on the other hand, is reinforced by a rival intuition to which one recent critic of the man-person distinction has appealed. My concern for my future is, at least normally, concern for what happens to *this* animal. I might indeed wish for psychological discontinuity if that is a means to this animal's future happiness. And if knowledge that this animal will suffer intense pain raises a fear for my own future (whose else?), it is a fear which does not seem rationally alleviated by the consideration that, before the pain occurs, my memories will be blotted out.¹³⁴ A neo-Lockean who wants to make a significant concession to this cogent point must here move to the view that persons are substantial, arguing perhaps that, although psychological continuity does not constitute personal continuity, it is not irrelevant to it in that the survival of the self is the survival of the seat of (normally continuous) consciousness. In the brain-transplant case, if the seat of my consciousness survived transplantation with undamaged basic function, what is important to me would survive, whether or not a consequence of operational shock was irreversible amnesia and alteration of personality.

There is, however, an obvious alternative to all forms of neo-Lockean explanation of the brain-transplant case, which is to deny that the person whose brain is supposed to

have been transferred really would continue to exist through a successful operation. Perhaps there would be a certain continuity important in prospect to that person. Perhaps too, since the being who emerges is supposed to enjoy the self-appearance of having existed as the brain-donor, it would be convenient to talk *as if* the person had survived the demise of the animal. But neither consideration would justify a judgement of identity.

Certain analogies reminiscent of Boyle's discussion of life may help to establish the reasonableness of this approach. First, let it be conceded for argument's sake that intelligence involving a capacity for a high degree of self-awareness together with memory of one's own past is a necessary and sufficient condition of being a person. We know that many members of at least one species of animal satisfy this condition for a large part of their lives, during which, at least, they are persons or, less portentously, people. Now some capacities of animals are seated in particular parts of the animal: e.g. the sense of sight in the eyes, optic nerves etc. If that part (or a significant part of that part) were surgically transferred from one animal to another, then it could be said, perhaps properly and certainly not unreasonably, that the capacity itself had been transferred, e.g. that William now has Jane's sight, or that Elizabeth has been given Peter's capacity to purify the blood or make sugar assimilable. Some noun-predicates among the class of 'compound terms' apply to things in virtue of specific capacities: e.g. 'musician', 'singer', 'stallion' and 'person'. In an earlier discussion of Boyle we imagined its being said, when a stallion lost its generative capacity, that the *horse* still exists, although the *stallion* is no more. Consider, then, the following thought-experiment: the gonads of a valuable stallion or brood-mare are transplanted, and the product of these gonads continues to have the genes of their original possessors. Could we not say of such a case, just a little wildly, that, although the *horse* is no more, the *stallion* or *brood-mare* continues to exist. What *matters* or is of value, the horse-breeders might say, and what determined that the original horse was *called* a stallion, and the sire of this or that offspring, is still with us. It follows (the more philosophical of the horse-breeders might add) that the horse and the stallion were really always different entities. A dispute might then break out as to whether in that case stallions and brood-mares are substances or modes; and if substances, whether they are co-material with horses or, rather, are parts of horses.

In so far as rationality, self-awareness and memory are linked capacities with their seat in the brain, then their surgical transfer from one rational animal to another would *exactly* parallel our imagined transfer of a stallion's generative power. (We can even imagine a stallion's 'splitting'.) Richard Overton, the seventeenth-century mortalist, saw the mind as the collection of faculties which raise man above the beasts, and he argued, not unreasonably, that no faculty or any other attribute can conceivably survive its subject. Transplant surgery shows that his conclusion is not unproblematically true even of animals, but nothing of ontological consequence follows from its falsity. From an ontological point of view the analogy between persons and stallions is complete. The difference between them, a difference which is responsible for the intractability of the problem of personal identity and the rich (but mappable) variety of responses which the problem has provoked, lies in the nature of the capacities in question. Like 'multiple personality', the transfer of someone's self-awareness would involve a kind of profound

illusion which for the purposes of talk or even thought it might in general be immensely convenient to fall in with. Yet the convenient assumption that the self-aware individual actually survives, taken with its obvious shortcomings, generates the tension which the neo-Lockean, at the cost of disastrous absurdity, seeks to alleviate by distinguishing the thing which does not survive from the thing which does, the rational animal from the person.

The neo-Lockean enterprise is powerfully aided and abetted by the thoroughgoing conceptualism of our time. Indeed, the support is mutual. Modern conceptualism itself has some of its longest and strongest roots in Locke's thought, not least in the chapter on identity. The *Essay* was perhaps the last great work of realism before the plunge into idealism. Yet it is a valuable object of study, not only as a source of alien and forgotten insights, but also for the light it casts on the origin of some of the pervasive dogmas of twentieth-century anti-realism.

Conclusions of Volume II

The volume just concluded falls into three parts, the first and last of which form a continuous argument about the category of substance and the claims of substances to logical and ontological primacy over non-substances or modes. The second part supplements the argument by exploring some of the distinctive features of Locke's division of the sciences between substances and modes, a division closely related to his dual conception of law. It attempted to explain the form of his philosophical mechanism, and his conception of natural science as an understanding of the laws which flow necessarily from the actual natures or essences of substances. It also examined the structure of his rationalist ethics and his theistic theory of obligation, themes which are closely linked to his theory of personal identity. The chief unifying theme of the volume has been the distinction and relationship between what is natural, real, independent or 'given' and what is notional, ideal, mind-dependent or constructed, a distinction which is most directly dealt with in the course of Parts I and III.

The various revisions of the traditional Aristotelian doctrine of substance which were proposed in the seventeenth century necessarily faced the problem of holding together in a new way the two halves of the theory which it was in some sense the very purpose of the 'new philosophy' to drive apart. For the Aristotelians the category of substance contains both the logically basic objects of ordinary discourse, sensible material things, and the fundamental objects of natural science, the things whose natures constitute the ultimate principles of explanatory theory. The category easily contains both these things just because, to put the Aristotelian view a little roughly, they are the same things. The message of the 'new philosophy', however, was that there is an epistemological gulf between the world as it is presented in experience, as stuffs and individuals classifiable into their kinds on the basis of their sensible qualities and testable powers, and the world as it needs to be conceived of for the purposes of natural science, as body or matter quantitatively modified and subject to intelligible mechanical laws. The mechanistic view of the material world also raised in a more insistent form the question of the place of consciousness or the mind.

The argument of the present volume identified something like the following problem for the 'new philosophy': crudely, if a substance is what is fundamentally there, the essence or nature of which is an ultimate object of science, then matter or spirit may be a substance but gold, horses and men are not substances as such; yet the latter have always been regarded as the paradigm substances, paradigmatically fulfilling those logical conditions of substancehood by which the category of substance has ordinarily been defined. It might therefore seem that, for the 'new philosophy', the traditional notion of substance breaks in half, so that matter is a substance in the more fundamental sense,

different from the sense in which a horse is a substance.

The most characteristic feature of Locke's sceptical theory of substance, however, his explanation of the logical status of the category and of the distinction between primitive noun-predicates and adjectival predicates, firmly anchored the notion of substance to the level of experience rather than the level of physical theory. The distinction exists, he claimed, just because of our ignorance of the essences or natures of the things which we can identify through their sensible qualities and effects. Roughly speaking, we have to think of a horse, gold and the like as *something* unknown which possesses those qualities and powers. Indeed, the analysis extends to the idea of matter or body itself. In forming even our best idea of matter we have not, after all, achieved an understanding of the ultimate object of physics, and the logical difference between *body* and *gold* is simply one of generality. Ideas of modes are different from ideas of substances just because they do not purport to be ideas of anything in the world. In all this Locke may be contrasted with Spinoza or Leibniz, for example, for whom the philosophical problem of substance was primarily that of hammering out a new way of conceiving of reality as it is in itself. For such philosophers the characteristic features of the category of substance reflect the way that the world must ultimately be, not our epistemic relation to it.

For Locke the distinction or contrast between substance and mode corresponds to the difference between the natural and the *a priori* sciences. A science of substances may be unattainable, but sciences of simple and mixed modes are capable of indefinite development because we are capable of understanding what we ourselves construct without reference beyond our own thought. Yet it has been seen that Locke did not always find the distinction firm and unproblematic. Although he was always ready to insist that substances have a unity which is independent of our thought, or at least not dependent on it in the same thorough-going way as the unity of non-substances, he also seems to have been inclined to hold that substances as we conceive of them are in some sense 'creatures of the mind', not only at the level of species, but also (in the second edition) at the level of the individual. Lockean persons in particular, despite his explicit view of them as substances, seem to lie uneasily on a borderline between substances and modes. For the kind of unity attributed to them, while no doubt supposed real enough, is taken to have significance for classification and individuation precisely because the *a priori* science of ethics requires the abstract concept of a rational being subject to law and liable to just punishment or reward.

Remarks as summary as these cannot sufficiently suggest the subtlety and intricacy of the logical and ontological debate to which Locke was so important and successful a contributor. The analysis of that debate offered in the present volume has supplied the context for a continuous argument in favour of taking traditional interpretations of the distinction between substance and non-substance more seriously than it is nowadays orthodox to take them. The only satisfying account of the conjunction of the logical properties of the category of substance, the only satisfying explanation of what have been called above 'compound' or 'composite' terms, and the only satisfying, unparadoxical theory of identity all demand that we recognize that substances or material objects characteristically have a kind of natural or given individuality, unity and continuity which non-substances cannot have. The individuality of substances is real, and prior to their

individuation by us. The noun-predicates which serve to say 'what they are' simply place such natural individuals in classes according to affinity, origin or structure. The individuality of non-substances, on the other hand, is posterior to our individuating them and to the predicates or concepts through which they are sliced out of reality. The exploration of this dichotomy, and a better explanation than Locke's of the peculiar unity of substances, make it possible to see that his quasi-conceptualist approach to identity, for all its originality and influence on later thought, is incoherent. In particular, any attempt to maintain a significant ontological distinction such as Locke's between the mass of matter and the living body, or the *person* and the *man*, is doomed to failure.

Yet all these conclusions go some way towards endorsing one aspect, at any rate, of Locke's theory of substance. One could say that one line of descent from the Aristotelian theory of substance lies in the direction notably taken by Leibniz, and is the sphere of those philosophical scientists and philosophers of science who aim to make sense of the world as it is dealt with by fundamental physical theory. But the other line of descent lies more in the way of Locke, and is directed towards a theory-neutral explanation of the famous logical features of the category of substance. As it has been argued above, these features reflect, not an outmoded way of theorizing about reality, but the contribution of reality itself, at the pre-theoretical level of ordinary experience and judgement, to the fundamental forms of our thought and language.

Notes

Note

References are primarily by page number to the editions listed under 'Bibliography', but in order to facilitate reference to some multiply edited and collected primary sources (or to make active reference unnecessary for those sufficiently familiar with the texts) passages are in some cases also identified by short title and/or chapter, section etc. In the case of references to Locke's *Essay*, the first line (only) of the passage is given after the page, thus: 'Locke 1975:104,13 (II.i.1)', i.e. page 104, line 13. When a reference is to the last work referred to, the page number (or volume and page number, or, in the case of the *Essay*, page and line number) is given without the name of the author or date of the edition.

Introduction

1 Cf. Wittgenstein 1958:14–18.

2 Cf. Quine 1965:18–24. Quine finds Wittgenstein's argument too wide in its scope, although otherwise comparable to his own (Quine 1969:31). Putnam 1981:52 expresses the general view: "'Objects" do not exist independently of conceptual schemes. We cut up the world into objects when we introduce one or another scheme of description.' Dummett 1973:563 and 577, approves the model suggested by Geach 1962, according to which the world is an 'amorphous lump' which 'can be sliced up into distinct objects in various different ways according to the direction of cut'. Put so bluntly, the view cannot perhaps be straightforwardly described as the current orthodoxy, in that many would share the dissatisfaction with it which seems to find expression in the arguments of Davidson and Strawson considered immediately below. Yet it is so characteristic of the philosophy of our time that few would-be rebuttals appear as more than qualms in the face of its implications: the conceptualist's premises are combined with a fastidious disapproval of the overwhelmingly natural conclusion from those premises. A more fundamental line of objection to conceptualism, however, has been opened up by the arguments of Kripke and Putnam, already mentioned in Volume I and discussed at various places below, with respect to natural kinds and, in Kripke's case, proper names. The 'conceptualist realism' of David Wiggins, considered in Part III below, is an attempt to draw out some of the implications of Kripke's views for the theory of identity, and so for ontology; but as Wiggins's self-labelling indicates, he does not attempt to throw off the general assumptions of conceptualism. Putnam's earlier inclinations towards realism seem, sadly, to have been completely smothered in his recent work (witness the quotation above). An attempt will be made below to explain how the Kripke-Putnam line of argument has involved only a part of the case for realism. Their theses need to be set in a wider theoretical context, and cannot in the end stand without the support, in some form, of other elements in the traditional theory of substance.

- 3 Ayer 1958:42.
- 4 Kuhn 1962.
- 5 Davidson 1984:183–98.
- 6 Cf. Strawson 1964 and 1966.
- 7 Locke 1975:148,14 (II.ix.13).

Part I: Substance and mode (pp. 13–128)

I Introduction to Part I

- 1 See Volume I, chapters 1 and 9 (logical form); 11 (method); 3, 4, 12, 18 and 21 (perception); 29 (innate ideas).
- 2 See Ayers 1977. The interpretive argument used in that piece differs from the argument below, although their general conclusions are similar. The former relies heavily on the relationship with Boyle, does not do much to place Locke in the logical tradition, and contains some errors of interpretation (most importantly, of the sentence at Locke 1975:295,21).
- 3 Locke 1975:10,4 (Epistle).
- 4 See, for example, 580,35 (IV.vi.4):

I have chose to explain this uncertainty of Propositions in this scholastick way, and have made use of the Terms of *Essences* and *Species*, on purpose to show the absurdity and inconvenience there is to think of them, as of any other sort of Realities, than barely abstract *Ideas* with Names to them.... Though therefore these things might, to People not possessed with scholastick Learning, be perhaps treated of in a better and clearer way: yet those wrong Notions of *Essences* and *Species*, having got root in most Peoples Minds, who have received any tincture from the Learning, which has prevailed in this part of the World, are to be discovered and removed, to make way for that use of Words, which should convey certainty with it.

The point may not have been very lucidly expressed, but it reveals Locke's clear-headed rhetorical purposes, and casts light on his similarly ambivalent attitude towards the substance-accident distinction. (Quite another construction is placed on this ambivalence in Bennett 1987.)

- 5 Malebranche 1980:242f. Perhaps this should read 'in order to understand the *terms*' (rather than '*things*'), since the original syntax is ambiguous.
- 6 Locke 1975:175,26 (II.xiii.20).

2Substance, essence and accidents before Locke

7 Whether or not Aristotle himself saw things in this way, his theory at least carried the seeds of this reductive understanding of a 'mode of being' of a substance. On the other hand some Scholastics argued that the distinction between substance and accident (such as quantity or quality) was not a mere distinction of reason, even if not a fully real distinction. Cf. Suarez 1866:255. Descartes, however, viewed accidents as mere modes of their substance. For a vigorous expression of the reductive view, cf. Boyle 1667:10 (1979:22):

When a Bowl runs along or lies still, that *Motion* or *Rest* or *Globose figure* of the *Bowl* is not *Nothing*, and yet...to make them *real* and *physical* Entities, (for we have not here to do either with *Logical* or *Metaphysical* ones) is, as if, because we may consider the same man sitting, standing, running, thirsty, hungry, weary, etc. we should make each of these a distinct Entity, as we do give some of them (as hunger, weariness, etc.) distinct names. Whereas the subject of all these Qualities is but the same Man as he is considered with [different] circumstances.

8 Aristotle 1984:II 1623–44 (1028a10–1041b32).

9 A less mundane exception was the Unmoved Mover, or God.

10 Spencer 1628:59f.

11 Spencer expressed this standard view: 'Properties be not adjuncts: for adjuncts do outwardly befall the subject: but, they are necessary emanations from the principles of nature.' (1628:62) An example of the less common rival view is Coke 1654:66f, where an accident is defined as 'that which sticketh to the subject; [and] is called also adjoynt', but properties are called 'proper accidents' as opposed to 'common accidents'. An account like Coke's had been advanced by Maimonides, according to Wolfson 1977:II 165.

12 This definition was the one Plato had arrived at by division. Cf. Locke 1975:500,5 (III.x.17). For Locke, Plato's definition is as good as Aristotle's, since all our definitions are nominal (and, as we shall see, all our kinds are compounds).

13 Aristotle 1984:II 1625 (1029a11). Aristotle's account of matter as the subject or substrate (*hypokeimenon*) of quantity was no doubt behind the explanation, given at Locke 1975:498,11 (III.x.15), as to how the traditional idea of *matter*, as distinct from the idea of *body*, came to be formed:

Matter and *Body*, stand for two different Conceptions, whereof the one is incomplete, and but a part of the other. For *Body* stands for a solid extended figured Substance, whereof *Matter* is but a partial and more confused Conception, it seeming to me to be used for the Substance and Solidity of *Body*, without taking in its Extension and Figure.

Locke claimed that the mistaking this arbitrary logical distinction for a real distinction is what has 'produced those obscure and unintelligible Discourses and Disputes, which have filled the Heads and Books of Philosophers concerning *Materia prima*'. See Alexander

1985:222f for another interpretation of this passage.

- 14 Cf. Aristotle 1984:I 525 (321a1). Whether the Scholastic notion of *materia prima* closely corresponds to anything in Aristotle's own catalogue of things in the world has in effect been the subject of recent debate. When matter is opposed to form in his works it is generally assumed that the matter is of a certain kind, as the matter of a house is bricks and mortar, the matter of bricks is clay and so forth, down to the elements, earth, air, fire and water. The crucial question is whether Aristotle ascribed a role to undifferentiated matter as the substrate of substantial change. In the end it seems that he did (cf., for example, Robinson 1974 and King 1956).
- 15 Descartes 1964–76:VII 84 (*Med.* VI).
- 16 Descartes 1964–76:28–32 (*Princ.* I 60–5).
- 17 Cf. VIIIA 42–7 (*Princ.* II 5–12). The distinction between a corporeal substance and its quantity Descartes attributed to a confused conception of the substance as incorporeal.
- 18 Cf. Locke 1975:295,20. Cf. also the definitions of 'accident' at Hobbes 1839–45:I 102–5 (*De Corp.* IV.xxv.2).
- 19 Descartes 1964–76:VII 161.
- 20 VII 271–3.
- 21 VII 338.
- 22 Locke 1975:175.1 (II.xiii.19).

3 Our complex ideas of substances and the idea of substance in general

- 23 Locke 1975:295,4 (II.xxiii.1f).
- 24 Leibniz 1981:218 (II.xxiii.2).
- 25 Locke 1975:305,4 (II.xxiii.14). This is yet another case in which Aristotelian theory and terminology is characterized as a vulgar error dressed up.
- 26 Digby 1645:2f. (Similar passages are cited in Volume I, chapter 3, of the present work.)
- 27 Hobbes 1651:207 (*Lev.* xxxiv). No doubt both Hobbes and Digby were influenced by the Second Meditation.
- 28 Locke 1975:119,5 (II.ii.1).
- 29 Locke 1990:1f (*Draft A* 1).
- 30 Locke 1990:129f (*Draft B* 19).
- 31 130.
- 32 Aristotle 1985:II 165 (99b35–100b5).
- 33 Locke 1823:IV 16f.
- 34 IV 17–19. Locke repeats his characterization of the argument of II.xxiii.1 in the footnote added *ad loc.* to the fifth edition of the *Essay*.
- 35 Cf. Locke 1975:308,11; 313,1 (II.xxiii.23 and 30).
- 36 Cf. 308,15; 310,10 (II.xxiii.23 and 26f). Joseph Glanvill had used a similar argument: 'If it be pretended for a difference, that the parts of solid bodies are held together by *hooks*,...I say,...the *coherence* of the parts of these *hooks*...will be as difficult a conception' (Glanvill 1661:50). But cf. note 54, below.
- 37 The question rests in part on the parsing of the opening sentence of II.xxiii.3. If we take the final expression, 'that Substance', to refer back, somewhat ungrammatically, to '[each of the]

particular sorts of Substances', the sentence will simply mean that, when we employ the notion of substance in general, we presuppose that it stands in for a different 'internal Constitution, or unknown essence' for each species defined. On the other hand, we might take 'that Substance' to refer back to 'Substance in general', a reading which requires us to suppose that Locke took 'in general' to go with 'Substance' rather than with 'Idea' (i.e. that the shift described in my next paragraph had already taken place in Locke's argument). We could then take the expression 'internal Constitution' to be alluding explicitly to the determinate arrangement, 'constitution' or 'modification' of determinable matter. Cf. 442,22 (II.xxiii.6): 'the real Essence is that Constitution of the parts of Matter, on which these Qualities, and their Union, depend'. Perhaps Locke intended the expression to have certain corpuscularian overtones, but nothing more definite at this point. At 417,17 (III.iii.15) he wrote of the 'artificial' or supposed 'constitution' of genus and species.

4 Substance and real essence, matter and spirit, and the obscurity and confusion of the idea of substance

38 Locke, 1975:300,10 (II.xxiii.8).

39 Cf. 300,8; 298,16; 298,18 (II.xxiii.6 and 8).

40 This characterization hardly fits the case of thinking things, of course, supposing them to be simple immaterial substances. Yet Locke sometimes (as in the account of personal identity) came very close to treating the postulated immaterial substance as the stuff of which thinking things might be wholly or partly composed.

41 Locke 1975:383,14 (II.xxxi.13).

42 Locke 1823:IV 82.

43 Cf. Locke 1924:233f and 154. Of the passage from the *Letter* Pringle-Pattison claims that it throws an instructive light on Locke's conception of 'naked substances to which any kind of qualities may be arbitrarily annexed'. For a brief bibliography of the exegetical controversy (to which Alexander 1985:204–35 and Bennett 1987 should be added) see Ayers 1977:78 (footnote).

44 *Mutatis mutandis*, given that the form of an axe works through the axe's maker, whereas the form of a natural substance was supposed to work naturally.

45 Essentially the same thought about the mutability of real specific essences in a corpuscularian world is expressed at Locke 1975:419,17 (III.iii.19), discussed below in chapter 6.

46 Locke 1823:IV 25.

47 Cf. Locke 1975:450,11 (III.vi.21). Here, in discussing whether the essence of body is extension, he characterized 'an extended solid thing' as the 'Essence...in respect of us', a formula which also embraces the nominal essence of man (450,19). The argument is discussed in chapter 5, below.

48 Cf. 543,10; 542,23; 541,5 and 15, etc.

49 Cudworth, in defending the possibility of animal souls, described the theological objection as 'the grand argument' against them, and countered that their natural indestructibility need be no more disturbing theologically than the natural indestructibility of matter (Cudworth 1743:745). Cf. Locke 1823:IV 460; Locke 1936:121 (Journal 20 February 1682).

50 Locke 1975:108,24 (II.i.10). Cf. 228,36 (II.xix.4).

- 51 Cf. Locke 1936:123: '[Spirit and matter] may both lye dead and unactive, i.e. the one without thought, the other without motion a minute an hour or to eternity, which wholly depends upon the will and good pleasure of the first author.'
- 52 Leibniz 1981:379.
- 53 Locke 1975:542,19. Cf. 541,13.
- 54 541,5. For further discussion of 'superaddition', see chapters 12 and 14, below. Locke's official, never retracted scepticism with respect to the issue of dualism deserves notice as a virtue of his argument by comparison with Gassendi's. Glanvill, another sceptic about essences, seems to have missed the connection altogether in an argument which otherwise echoes Gassendi's and anticipates Locke's. For Glanvill, 'the soul is seen, as other things, in the *Mirrou* of its effects, and attributes: But, if like children they'll run behind the glass to see its *naked face*, their expectation will meet with nothing but vacuity and emptiness.' Yet at the same time he claims dogmatically that soul and matter are evidently 'distant in extreams', differing 'in their essentiall compositions' (Glanvill 1661:19–21).
- 55 Locke 1823:IV 33–7.
- 56 i.e. at 305,19 and 26. 'Immaterial' occurred in the first edition at 305,22. Similar changes were made in later paragraphs (at 306,16; 307,26 and 29; 310,31; 313,16; 314,5 and 21) again reinforcing first edition employment of 'immaterial' (cf. 308,1 and 314,13). The argument itself was unaffected, since from the first Locke's explicit concern was with our ideas of body and spirit 'as contradistinguished' (306,17), and with the conceptual possibility 'that Thinking should exist, separate, and independent from Solidity' (314,7).
- 57 306,3 (fourth edition).
- 58 Locke 1823:IV 33. Cf. Locke 1976–88:IV 600 (Molyneux to Locke, 22 December 1692).
- 59 Another place at which Locke speaks of probability in this connection is at 1975:345,25 (II.xxvii.25), where the wider context is concerned with immortality. He may have felt that (despite his own sceptical arguments) the possibility of immortality is better catered for by dualism than by materialism: so that, given the probable truth of Christian revelation, dualism is that much more probable than materialism. But 'that much', by his own reckoning, should be rather little.
- 60 Cf. Cudworth 1743:755, where it is argued that the 'resolution' of secondary qualities into 'mechanism and fancy' presupposes sensation and imagination, faculties whose operation cannot be similarly resolved: 'wherefore undoubtedly they are no modes of matter or body, but attributes of another kind of substance incorporeal'.
- 61 Cf. Bennett 1987:209.
- 62 Locke 1975:174,4 and 7. See Volume I, chapter 25 for discussion of Locke's account of space.
- 63 363,18 (II.xxix.2).
- 64 365,5 (II.xxix.7).
- 65 175,10 (II.xiii.19).
- 66 The *quaestio* in question might be 'Is space something or nothing?' Locke, following Gassendi, was attacking the maxim standardly brought to bear on this question by the nihilists, 'Everything is either substance or accident'.
- 67 Cf. Leibniz 1981:150,218 and 305 (II.xiii.19f, II.xxiii.2, III.vi.4).

5 Substance, mode and the argument from language

- 68 Locke 1975:297,4.
 69 450,5 (III.vi.21).
 70 450,21. As to who might have been taken to have claimed that rationality is the *whole* essence of man see note 79, below.
 71 Descartes 1964–76:I 30f (*Princ.* I 63).
 72 Malebranche 1980:244.
 73 Locke 1975:493,29 (III.x.6).
 74 Malebranche 1980:243.
 75 245. It is significant that the version of this argument given in Malebranche's *Dialogues on Metaphysics* III is not directed against Aristotelian matter, but against the suggestion (equivalent to Gassendi's assertion that we are ignorant of the essence of matter) that the divine archetype of matter is as unknown to us as is (according to Malebranche) the archetype of spirit (Malebranche 1923:109f). But the *Dialogues on Metaphysics* were published in 1688, too late to have influenced the form of Locke's appeals to language.
 76 Some Scholastics did, however, hold that it was within the power of God to maintain matter in existence without form, and Malebranche may have had this claim in mind. Cf. Suarez 1866:257.
 77 Locke 1823:IV 19. In fact Locke had already in 1671 assimilated the notions of 'being' and 'something', but had understood them in a way which had rather different implications with respect to our ideas of substances. 'Being' or 'something', he then held, is just a name we use for everything we can think of. Consequently our notion of *being* comprehends, but cannot extend beyond the simple ideas we get from experience. On this view the inclusion of an idea of being in our ideas of substances would add nothing: 'all the notion we have of substance amounting at last to noe more than the Ideas of certain powers', Locke 1990:20f (*Draft A* 8). This treatment of the idea of existence was dropped from *Draft B*, without replacement. Had it been retained in the *Essay*, one of Stillingfleet's objections would have been more appropriate than it in fact was, i.e. his attributing to Locke the absurd view that the idea of substance in general is a complication of many ideas together (cf. Locke 1823:IV 16).
 78 Arnauld and Nicole 1965:46–9.
 79 That Locke, at III.vi.21, took the trouble to attack the untraditional suggestion that rationality (*ratio*) is the whole essence of man (rather than *rational animal*—cf. note 70, above) suggests that he had the present passage from the *Logic* there in mind; and therefore also suggests that he had it in mind at III.viii.2 too, as offering a Cartesian explanation of the Scholastics' solecism for which he needed to offer an anti-dogmatic alternative.
 80 Locke 1975:475,10 (III.viii.2). In other words, 'humanity' is more closely linked to 'humane' than to 'man'. I should have liked to have included an account of the argument, also appealing to the 'ordinary Words of Language and our common use of them', of III.viii.1; but Locke's thought here has finally remained for me too slippery for cost-effective discussion. It might be that the right comparisons with Scholastic discussions of abstract terms would make determinate interpretation possible; or it might be that the thought was never more than indeterminate and confused. I commend the section to the consideration of others as one of the darkest places in the *Essay*.
 81 There are more interesting counter-examples: for example, 'is a triangle' is as satisfactory a

predicate as 'is triangular'.

- 82 Cf., for example, Locke 1975:436,27 (III.v.14). Locke was probably adapting Hobbes here: 'Geometry therefore is demonstrable, for the lines and figures from which we reason are drawn and described by ourselves; and civil philosophy is demonstrable, because we make the commonwealth ourselves' (Hobbes 1839–45:VII 184). For Locke, what we relevantly make is not the thing, but the idea.
- 83 585,20 (IV.vi.11).
- 84 123,26 and 27 (II.iv.1), quoted with my emphasis.
- 85 Coste's note on IV.x.18 (Locke 1729:xlivf). The speculation perhaps owed something to Descartes' theory of the constant recreation of matter and motion.
- 86 Locke 1975:329,25 (II.xxvii.2).
- 87 Possibly a distinction of meaning or sense between the predicative copula and the 'is' of identity would not figure in the best ultimate account of these differences, but the best ultimate account is not necessary for present purposes. The distinctions drawn here are more or less ignored in recent philosophy and yet obviously call out for serious explanation.
- 88 Leibniz 1981:218.
- 89 Duncan 1748:28–42.
- 90 Kant 1962:97 (*Prolegomena*, sect. 46). It is worth reading the whole of this section for its heavily Lockean overtones.

6 *Species and their names in the corpuscularian world*

- 91 Cf. Sloan 1972:1–52. Although my reading of Locke's argument differs from Sloan's, it is consonant with his persuasive historical thesis.
- 92 Locke 1975:410,33 (III.iii.6–11). Cf. Descartes 1964–76:VIII 27f (*Princ.* 58f); Hobbes 1839–45:I 23f (*De Corp.* I.ii.14).
- 93 Locke 1975:460,10 (III.vi.32); 417,16 (III.iii.15).
- 94 Together (in Descartes' case) with a premise about God's nature and activity. See chapter 11, below.
- 95 Hence the capacity to shock of the radical suggestion, at Locke 1975:308 (II.xxiii.23), that there is something occult even in extension.
- 96 The question whether Aristotle himself had really advanced the doctrine of substantial forms at all, since he was so fond of illustrating the matter/form distinction with the example of a bronze sphere, is provocatively raised at Boyle 1667:72 (1979:53). Cf. 42f (1979:39).
- 97 It may seem odd to compare metals and chemicals with machines, but cf. Boyle 1667:54 (1979:45).
- 98 Cf. Locke 1975:580,3 and 470,9 (IV.vi.4 and III.vi.50).
- 99 418,11 and 415,15 (III.iii.17 and 13).
- 100 Cf. 462,14 (III.vi.36): 'Nature makes many particular things, which do agree one with another, in many sensible Qualities, and probably too, in their internal frame and Constitution.'
- 101 447,10 (III.vi.12); cf. 451,8 (III.vi.22). For the influence of this understanding of the 'great chain' on eighteenth-century biology, see Lovejoy 1960:227–41. Lovejoy wrongly assumed that Locke's talk of real essences implies a belief in unknown real species, but he also

recognized that Locke's conceptualism could point in the same direction as Bonnet's 'If there are no cleavages in nature, it is evident that our classifications are not hers', or as Buffon's 'In general, the more one increases the number of one's divisions in the case of the products of nature, the nearer one comes to the truth; since in reality individuals alone exist in nature' (230f).

- 102 Locke 1823:IV 91 (discussed in chapter 4, above).
 103 Locke 1975:419,18 (III.iii.19). Cf. his explanation of the eternal truths at 638,20 (IV.xi.14).
 104 416,5 (III.iii.13).
 105 463,25 (III.vi.39).
 106 Such passages occur at 418,9 (III.iii.17); 448,20 (III.vi.14–19); 452,24 (III.vi.25). In each of these the epistemological objection is brought after ontological objections have been concessively set aside. In the last, it is argued that, even if there *were* philosophers capable of discovering real essences, 'they have not been Philosophers, or Logicians, or such who have troubled themselves about *Forms* and *Essences*, that have made the general Names'.
 107 For Locke's commentary on this exception, see 464,23 (III.vi.40).
 108 Cf. 379f (II.xxxi.6), *et passim*.
 109 449,15 (III.vi.19).
 110 Cf. 449,22–8.
 111 Cf. 503,28 (III.x.22).
 112 483,32 (III.ix.14).
 113 486,23 (cf. 486,11) (III.ix.17).
 114 417,6 (III.iii.15).
 115 501,25 (III.x.19).
 116 441,37 (III.vi.5).
 117 440,18 (III.vi.4).
 118 442,15 (III.vi.6). For 'particular' in this common enough use cf. 296,24 (II.xxiii.3): 'the particular internal Constitution, or unknown Essence of that Substance', where the substances in question are 'particular sorts of Substances', not individuals. For the general argument and similar allusion to the predicables, cf. Boyle 1667:43 (1979:39):

Nor need we think that Qualities being but Accidents, they cannot be *essential* to a Natural body: for Accident...is sometimes oppos'd to Substance, and sometimes to Essence: and though an Accident cannot but be accidental to Matter, as it is a Substantial thing, yet it may be essential to this or that particular body...since an aggregate or convention of Qualities is enough to make the portion of Matter tis found in what it is, and denominate it of this or that determinate sort of Bodies.

- 119 For knowledge of these antecedents I rely on memories of an interesting talk by J.Needham, on what he called 'aurifaction' and 'aurifiction' in ancient China.
 120 Cf. Locke 1975:461,24 (III.vi.35): 'Should there be a Body found, having all the other Qualities of Gold, except Malleableness, 'twould, no doubt, be made a question whether it were Gold or no.' Locke's view is that the answer depends strictly on our nominal essence, and that in fact we are prepared to allow non-malleable gold, since 'eager' gold 'will as little endure the Hammer as Glass it self'.
 121 Cf. Boyle 1667:62 (1979:49): there is

a vast multitude of Portions of Matter endowed with store enough of differing Qualities to deserve distinct Appellations; though for want of heedfulness and fit words men have not yet taken so much notice of their less obvious Varieties, as to sort them as they deserve and give them distinct and proper Names.

But for the ultimate arbitrariness of the business, cf. 1667:39 (1979:37):

Men having taken notice that certain conspicuous Accidents were to be found associated in some Bodies, and other conventions of Accidents in other bodies, they did for conveniency and for the more expeditious expression of their Conceptions agree to distinguish them into several sorts.

- 122 See especially *Essay* II.xxix and III.ix–xi. The issue is discussed in Volume I, chapter 30.
 123 Locke 1975:456,17 and 33 (III.vi.29).
 124 Cf. 365,5 (II.xxix.7); 521,19 (III.xi.24); 648,20 (IV.xii.14) etc.
 125 Cf. 407,10 (III.ii.5); 378,20 (II.xxxi.6); 499,27 (III.x.17–21) etc.
 126 484,31 (III.ix.16).
 127 456,16 (III.vi.29).
 128 522,15 (III.xi.25).
 129 521,6 (III.xi.24).
 130 378,25 (II.xxix.6).
 131 Cf. 443,18; 447,30; 463,14 (III.vi.8, 13 and 39).
 132 520,9 (III.xi.22). Cf. 488,10 (III.ix.21) and 500,19 (III.x.18).
 133 378,12 (II.xxxi.6–11).

7 Are there real species?

- 134 The following few pages owe an obvious debt to arguments in Kripke 1972 and Putnam 1975, although the issue of ‘natural kinds’ is here conceived of rather differently (cf. note 137, below). For Leibniz’s defence of a partly quasi-Aristotelian position against Locke’s attack, see Leibniz 1981:308ff (well discussed by Jolley 1984:145–61).
 135 Locke 1975:381,29 (II.xxxi.9). It is in fact a bit odd that Locke should have suggested, in this somewhat maverick passage, that a rational observer would be impelled *at first glance* to leave size and shape out of any nominal essence formed. The suggestion seems neither consonant with his general account nor even approximately true. What if the lump of gold was not only the observer’s first observed gold, but also their first observed statue or artifact?
 136 577,16 (IV.v.7).
 137 Cf. Putnam 1975:225 and 229f, or Kripke 1972:31: ‘The original concept of cat is: *that kind of thing*, where the kind can be identified by paradigmatic instances.’ Kripke is accurately characterizing both Putnam’s view and his own. Neither seems even aware of the traditional question how the boundary of a ‘kind’ is determined, or how we determine it.
 138 Walters 1961:83.

- 139 Ford 1975:70.
 140 Locke 1975:452,6 (III.vi.23).
 141 464,15 (III.vi.39).
 142 416,5 (III.iii.13).
 143 Boyle 1979:44.
 144 Locke 1975:440,24 (III.vi.4).
 145 418,16 (III.iii.17).
 146 As we explain the distinction between white and yellow gold. But contrast Boyle 1979:39:

for the most part a concurrence of many [qualities] is so essential to [each] sort of bodies that the want of any of them is sufficient to exclude [a body] from belonging to that species...as the chemists' *luna fixa*, which they tell us wants not...any other property of gold except yellowness (which makes them call it *white gold*), would by reason of that want of colour be easily known from true gold.

- 147 Locke 1975:419,21 (III.iii.19).

8 Locke on the difference between substances and modes

- 148 Locke's medical activities no doubt at a general level confirmed his 'empiricism' (as well as his avoidance of the word 'empirical'), but had little more specific effect on the *Essay* except with respect to botanical classification, via an interest in vegetable remedies.
- 149 This is said in full awareness of the view that the scientist's is just one practical interest among many (say, an interest in predicting). It seems preferable to believe that the language of science is geared to telling a story which at least approximates to the ultimate nature of things. That is why it is an aid to prediction. Of course, it can be claimed that 'the ultimate nature of things' is just an honorific name we give to what matters for the important purpose of predicting, and that things have no nature (indeed, there are no 'things') independent of how they are conceived of for some human purpose or other. The argument of the present work is intended to demonstrate that such easy pragmatism is quite incapable of dealing with the logical explicanda.
- 150 Cf. e.g. Locke 1975:290,26 (II.xxii.6).
 151 Cf. 294,18 (II.xxii.11).
 152 Cf. 474,33 (III.viii.2): 'some concrete Names of Relations, amongst Men chiefly, are Substantives; as *Paternitas, Pater*; whereof it were easy to render a Reason'. That 'father' is really the name of a relation rather than of 'the thing [the relation] is attributed to' is argued at 322,12 (II.xxv.8). See above, chapter 5.
- 153 These issues relating to artifacts and persons will be approached through the topic of identity in Part III, below.
- 154 Cf. Mackie 1976:99f.
 155 Locke 1975:463,14 (III.vi.39). Cf. 331,21 (II.xxvii.5).
 156 Boyle 1667:342.
 157 Cf. Locke 1975:318,21 (II.xxiv.3).
 158 320,36 (II.xxv.4); cf. 322,12 (II.xxv.8). Yet at 360,32 (II.xxviii.18), in trying to illustrate the principle that '*all Relation terminates in...simple Ideas...* got from Sensation, or

Reflection', Locke seems to say just the opposite: the idea of a *friend* contains 'all the simple *Ideas*, comprehended in the Word Man, or intelligent Being', and that of a *father* contains the idea of 'that particular Species...signified by the Word Man'. In the case of *friend* it would seem possible to avoid a contradiction, and explain how it is that not every different nominal definition of *man* will give rise to a different idea of a *friend*, by arguing that the idea of man incorporated in the idea of a friend is the idea of the mode-like 'moral man' discussed at 516,31 (III.xi.17): hence a rational parrot or monkey could be a friend. Even so, this seems to be a different explanation from the one given in II.xxv, and does not easily fit relations such as *father*, distinguished by Locke himself, at 349,18 (II.xxv.2), as 'natural relations'. In this last passage too, however, the suggestion is made that terms like '*Father and Son, Brothers, Cousin-Germans, etc.*' are restricted to our own species ('tis seldom said, This Bull is the Grandfather of such a Calf; or that two Pidgeons are Cousin-Germans') just because of their forensic employment for relations from 'whence...arise the Obligations of several Duties amongst Men'. There seems little doubt as to the general line of thought consistently pursued by Locke, but he was (of course) defeated in the attempt to treat *father* as both a 'natural' concept and a moral one.

159 Cf. 326,15 (II.xxvi.4).

160 447,30 (III.vi.13).

161 613,14 (IV.viii.6).

162 386,18 (II.xxxii.7).

163 293,13 (II.xxii.10). 'Laws' include, of course, the Moral Law.

164 361,26 (II.xxviii.26).

165 Cf. Volume I, chapter 30.

166 Cf. 429,11 (III.v.3–6); 288,14 (II.xxii.2–5).

167 436,19 (III.v.13). For one place at which he does discuss natural non-substances, see note 158, above.

168 431,18 (III.v.6). Cf. 290,8 (II.xxii.5–7).

169 434,34 (III.v.10). Cf. 289,25 (II.xxii.4); 429,18 (III.v.3) etc.

170 i.e. in cases where the doubt is not due to factual ignorance of (say) the physiology of the doubtful specimens: where the question is *purely* one of classification.

171 Locke 1975:291,21 (II.xxii.8).

172 329,29 (II.xxvii.2).

173 383,19 (II.xxxi.13).

174 376,33 (II.xxxi.3).

175 469,7 (III.vi.47). For the significance of Adam, cf. Aarsleff 1982:25–7.

176 Cf. Dummett 1973:73–80. Dummett associates the distinction with Frege and Wittgenstein, and a similar distinction is central to the thought of Quine.

9 Reflections on the notion of substance

177 See Volume I, chapter 15.

178 This may seem surprising, given the popularity of the notion of immaterial substance, but Aristotelian philosophy had to make a special exception to allow the existence of disembodied souls as form without matter (as Christian mortalists often pointed out with relish). Hobbes, as we have seen, took 'immaterial substance' to be a contradiction in terms.

- 179 Cf. chapter 2, above. The widely employed principle that substances are those entities which can be conceived or understood in and by themselves is not separately listed. The thought is variously covered by the first, second and fifth principles.
- 180 The most abstract and decisive difficulties are perhaps as follows. Immaterial substances would have either to exist in space or out of it. If the latter, then it is impossible to account for their individuality: we are lumbered with the principle of the identity of indiscernibles in its unacceptable, pure Leibnizian form. Numerical identity becomes comprehensive qualitative identity, so that individual immaterial substances become, in effect, complex universals. If, on the other hand, immaterial substances are supposed to exist in space, it becomes impossible to account for their place-occupancy. For if we allow them the kind of interaction with their environment which goes with intelligible place-occupancy, there is no ground for distinguishing them from other place-occupants as non-physical or non-material. To make them very small and like electrons might make them in a sense non-material, but not non-physical. Yet, in any case, why should they be supposed small and simple rather than large and complex. Their internal complexity as minds or thinking things must surely be reflected in the complexity of the causal interactions associated with their space-occupancy (as is the case with brains). Anything capable of complex thought must be an appropriately complex space-occupier. Of course, there are other problems with dualism, but these will do for a start.
- 181 Nagel 1979:165–80.
- 182 Frege 1968:42.
- 183 500,36 (III.x.19). Locke's claim is that, although this traditional point about non-substances ought to be extended to substances, the prejudice that there are natural kinds of substances has prevented people from doing so. The claim is, of course, doubly implausible, at least in cold blood.
- 184 Wiggins 1967:31ff, 59f and 71. Cf. Wiggins 1980:63ff.
- 185 See note 180, above.
- 186 Harris 1751:38.
- 187 Wiggins 1980:25f. Cf. also Mellor 1985:119–39.
- 188 Cf. Locke 1975:294,1 (II.xxii.11): '*Power* being the Source from whence all Action proceeds, the Substances wherein these Powers are, when they exert this Power into Act, are called *Causes*.'
- 189 There is a familiar question as to how far it is ever possible to identify entities postulated by one theory with those postulated by a different theory; e.g. Einstein's atoms with Newton's. If it is admitted as possible, then it must be supposed possible to identify or individuate theoretical objects while having only an approximate knowledge of their nature. Nevertheless, identifying an object through an approximately correct theory is different from picking out, without knowing what it is, a strange material object.
- 190 i.e., in the chapter on identity and diversity added to the second edition as II.xxvii.

Part II: God, Nature and the Law of Nature (pp. 129–202)

10 Introduction to Pan II

- 1 Locke 1958:108f.
- 2 I should here record a debt to a lucid treatment of the theological background in an unpublished (1989) typescript on 'Locke's Moral Theology' by Ian Harris. Moreover, despite our different interests in Locke's writings and different judgements as to their significance, it was Harris's forceful and original argument which persuaded me to put together discussions of causal and moral law kept separate in an earlier version of the present work.
- 3 Cf. Copleston 1962:III (Part I) 120.
- 4 Locke 1975:560,3 (IV.iii.29).

11 Forms of mechanism before Locke

- 5 'Purity' could of course be measured in different ways in this context: it is enough if the present usage is neither obscure nor arbitrary.
- 6 Arnauld and Nicole 1965:43. But cf. 167: 'It must be God who has impressed motion on matter and who maintains it there.'
- 7 Descartes 1964–76:VIII A 62ff (*Princ.* II 37).
- 8 Cf. Gabbey 1980:245. The translation of Beeckman has been slightly amended.
- 9 Cf. Gueroult 1980:212. Contrast Hatfield 1979:129. Hatfield cites Westfall and Carteron as holding similar views to Gueroult's on this point.
- 10 Hobbes 1839–45:131 and 213 (*De Corp.* II.x.6 and III.xv.3).
- 11 Descartes 1964–76:VIII A 66 (*Princ.* II 43).
- 12 Cf. VI 42ff (*Disc.* V).
- 13 Cf. Gabbey 1980:265,313.
- 14 Cf. Descartes 1964–76: VIII A 55 (*Princ.* II 27).
- 15 Gueroult 1980:198f; Descartes 1964–76:V 403f.
- 16 Following Hatfield 1979.
- 17 For a recent discussion of the relation between Descartes' and Malebranche's positions (especially the problem of differentiating between the doctrines of constant recreation and occasionalism) cf. McCracken 1983:89–115.
- 18 Boyle 1685–6:399 (cf.125,127); Boyle 1979:167 (*An Essay concerning those that would exclude the Deity from intermeddling with Matter*).
- 19 Boyle 1744:III 516f (*Reconcilableness of Reason and Religion*).
- 20 Malebranche, who laid considerable stress on the notion of God's ordinary and general concurrence, was only unusual in recognizing that the reconciliation had failed. The phrases

quoted come from Boyle 1979:139 and Boyle 1685–6:10.

- 21 389, 285 and 397. But cf. Boyle 1744:V 520 (*Christian Virtuoso*) for what seems a plain statement of the inertness of matter: ‘the laws of motion, without which the present state and course of things could not be maintained, did not necessarily spring from the nature of matter, but depended upon the will of the divine author of things’.
- 22 Boyle 1744:III 434 (*Excellency of Theology*). Cf. Boyle 1685–6:18ff.
- 23 242.
- 24 279.
- 25 Boyle 1979:161. Cf. 156f, where Boyle seemingly envisaged the possibility that gravity should be among phenomena due to ‘certain primitive, general, and fixed laws of nature’, rather than a case of ‘the admirable conspiring of the several parts of the universe to the production of particular effects’.
- 26 Bentley 1958.

12 *The form of Locke’s mechanism*

- 27 Locke 1975:540,24; 541,20.
- 28 544,21 (IV.iii.10).
- 29 544,30 (IV.iii.11).
- 30 545,11 (IV.iii.12).
- 31 545,16 (IV.iii.13).
- 32 556,5 (IV.iii.25).
- 33 558,29 (IV.iii.28).
- 34 379,7 (II.xxxi.6). Discussed by Wilson 1979.
- 35 Leibniz 1981:379–83. Compare Leibniz’s letter to Hartsoeker, criticizing the Newtonians as he understood them: ‘Thus the ancients and moderns who avow that gravity is an *occult quality*, are right if they mean thereby that there is a certain mechanism unknown to them, by which bodies are impelled toward the centre of the earth. But if their notion is that this transpires without any mechanism, by a simple *primitive property*, or by a law of God which brings about this effect without any intelligible means, then it is a senseless occult quality, which is to say that it can never be cleared up, even though a Spirit, not to say God himself, were endeavouring to explain it’ (cited by Cajori at Newton 1962:II 668). Newton’s editor Cotes responded indignantly to Leibniz’s objections to gravity: ‘shall gravity therefore be called an occult cause, and thrown out of philosophy, because the cause of gravity is occult and not yet discovered?’ (I.xxvii). It is difficult not to see Leibniz’s comments on Locke as part of the same polemical pattern as his criticisms of Newton.
- 36 McCann 1983.
- 37 Locke 1975:559,25.
- 38 546,11.
- 39 Both points are, indeed, explicitly alluded to in this section, at 546,29 and 17 respectively (not to speak of 545,35).
- 40 Cf. Locke 1823:IV 460f. This passage, and the meaning of ‘superaddition’, are further discussed in chapter 14, below.
- 41 Locke 1975:559,17 (IV.iii.28).
- 42 544,31 (IV.iii.11).

- 43 Cf. 554,19 (IV.iii.23).
 44 560,14.
 45 545,18 (IV.iii.13).
 46 546,12.
 47 123,17 and 27; 126,13 (II.iv.1, 2 and 5).
 48 541,34 (IV.iii.6).
 49 Locke 1823:IV 462ff. Cf. Boyle's category of the 'inexplicable': 'though we cannot deny *that they are*, yet we cannot clearly and satisfactorily conceive *how they can be* such as we acknowledge they are' (Boyle 1979:211).
 50 We might compare his procedure here with that adopted in putting forward problem cases in his argument about personal identity at Locke 1975:347,23 (II.xxvii.27). They are assumed to be possible although, with greater knowledge, 'we might see the Absurdity of some of these Suppositions I have made'. Possibility, necessity and impossibility have to be judged according to our ideas.
 51 Boyle 1685–6:18.
 52 Locke 1989:246 (sect. 192). Yolton's note assumes Locke's agreement with the view that God is the only real cause.
 53 Locke 1975:135,31 (II.viii.11) and textual footnotes.
 54 Locke 1823:IV 467f.
 55 III 304f.
 56 Locke 1975:547,29 (IV.iii.16).

13 Reflections on rationalism, empiricism and mechanism

- 57 This interpretation of Hume as a 'sceptical realist' is advanced in Wright 1983. Hume everywhere states that powers are unintelligible to us, and does so on the basis of a rationalist account of what it would be to understand a power. Cf. Hume 1958:161 (*Treat.* I.iii.14):

We must distinctly and particularly conceive the connexion betwixt cause and effect, and be able to pronounce, from a simple view of the one, that it must be follow'd or preceded by the other. This is the true manner of conceiving a particular power in a particular body.... Now nothing is more evident, than that the human mind cannot form such an idea of two objects, as to conceive any connexion betwixt them.

Hume's starting-point, at least, is thoroughly Lockean. Cf. G.Strawson 1989.

- 58 Locke 1975:235,5 (II.xxi.4). The necessity of the principle of universal causation is forcefully stated at 622,28 (IV.x.8).
 59 Russell 1959:67. The phrase 'under the same circumstances' is grammatically odd, but presumably means 'given regular association without exception'.
 60 Other paradoxes for inductivism have been remarked: for example, it is clear that observation of an instance of a low-level generalization can actually reduce the probability of the generalization. The hypothesis that no land-animal could exist weighing more than ten tons would be weakened, rather than strengthened, by the discovery of an animal thriving at a

pound or two under ten tons. Rationalism has no problem with such a case.

61 Cf. Strawson 1963:245ff:

The evidence for a generalization to the effect that all *As* are *Bs* is good (1) in proportion as our observations of instances of *As* which are *B* are numerous, and (2) in proportion as the variety of conditions in which the instances are found is wide; always provided that no instance is found of an *A* which is not *B*...But the need for condition (2) does not imply the existence of two fundamentally contrasted kinds of inductive reasoning. Suppose we frame a higher-order generalization to the effect that all generalizations of the form 'All *A*—s are *Bs*' (e.g. 'All *ACs* are *Bs*', 'All *ADs* are *Bs*' etc.) are true. Then the truth of 'All *ACs* are *Bs*' constitutes one favourable instance of this generalization, the truth of 'All *ADs* are *Bs*' constitutes another favourable instance and so on. By condition (1), the greater the number of such generalizations we establish, the better the evidence for the higher-order generalization. But the higher-order generalization and the original generalization ('All *As* are *Bs*') have exactly the same force, are logically equivalent.

That something is wrong with this argument is clear from the point that there is no determinate principle of counting subordinate generalizations: any range might be subdivided indefinitely according to a variety of incommensurable principles. Where there seems to be one obviously preferable principle (e.g. there is a determinate number of chemical elements), that is where nature obliges, and has nothing to do with logic. The significance of a variety of conditions can hardly be restricted to such cases.

62 Descartes 1964–76:IXB 327f (*Princ.* IV 205).

63 Hume's quasi-Baconian discussion of directed experiment at *Treat.* I.iv.15 (Hume 1968:173ff) presupposes that the principle of the uniformity of nature (in a particularly strong form: 'The same cause always produces the same effect, and the same effect never arises but from the same cause') is a principle which 'we derive from experience'. His general theory, however, commits him to the view that there is nothing intrinsically reasonable about putting hypotheses under stress.

64 Discussed in Volume I (chapter 25) of the present work.

65 Cf. chapter 7, above.

66 Empiricists have perhaps been more fond of the related argument that, since logical necessity is a relation between propositions or sentences, while nomic or causal relations hold between events, it is mere and inexcusable confusion to assimilate the two.

67 There is, of course, much more to be said on this topic: e.g. such propositions as 'The stuff Smith's ring is made of is soluble in *aqua regia*' or 'The particle discovered by X has negative mass' raise further issues. Moreover, the present argument is not concerned with causal claims in general, even though every causal claim implies the existence of a nomic relationship or relationships such as have been considered.

68 Hume 1958:230f (*Treat.* I.iv.4).

14 The existence of God

- 69 Locke 1823: IV 461.
- 70 Cf. Ayers 1981:231f, where its employment by John Norris is examined.
- 71 Cudworth published in 1678, and Locke was reading the work on 18 February 1682 (Locke 1936:118). Von Leyden 1948 draws attention to similarities to Nicole's less complex argument in *Essais de Morale*, which Locke translated (Locke 1712). Although Locke was evidently following Cudworth, there are some echoes of the translation in his argument. Von Leyden notices that Nicole's argument is unlike Locke's in embodying a proof of the soul's natural immortality, but he does not draw out the implications of this difference with respect to the main argument.
- 72 Cudworth 1743:667f, 684ff and 727–62.
- 73 862.
- 74 758.
- 75 862. Cf. 742.
- 76 755.
- 77 862.
- 78 750.
- 79 For relevant discussions of miracles and personal identity see Volume I (chapter 14) and chapter 23, below.
- 80 Locke 1823:IV 460f.
- 81 Cudworth 1743:745. The theological objection is characterized as 'the grand argument' against animal souls, to which Cudworth responds that their natural indestructibility is as theologically innocuous as the indestructibility of matter or any substance *qua* substance. Locke's journal entries of 18 and 20 February 1682 indicate that Cudworth's discussion stimulated his first important extant comments on personal identity. See chapters 20 and 21, below.
- 82 Locke 1823:466 (Second *Reply*).
- 83 Locke 1975:624,26 (IV.x.10).
- 84 623,10 and 27; 624,7 (IV.x.10).
- 85 Cudworth 1743:758. Cf. Locke 1975:623,20.
- 86 Cudworth 1743:758f.
- 87 Cf. Locke 1975:625,11 (IV.x.12).
- 88 Descartes 1964–76:VII 103f (First *Reply*). Locke was perhaps influenced by a parenthetical argument of Cudworth's which does without the dualist presuppositions of the main argument: even if everything created were made out of pre-existing matter, a further efficient (no doubt, intelligent) cause would be necessary since 'marble cannot make a statue, nor timber and stones a house, nor cloth a garment' (Cudworth 1743:758).
- 89 Locke 1975:621,11 and 16; 622,10 (IV.x.6f). Intimations of a quite general argument from design also appear at 624,24 and 625,7 (IV.x.10 and 12).
- 90 623,10 (IV.x.10).
- 91 623,29.
- 92 624,7 (IV.x.10). The final sentence was added to the second edition.
- 93 624,2.
- 94 627,10 (IV.x.16). There are strong echoes of Locke's translation of Nicole here, but

significantly Nicole was arguing at this point for the immateriality and natural immortality of the soul. The verbal similarity only emphasizes the different context of Locke's sentences in the *Essay*. For Locke, an argument like Nicole's only works with respect to *God's* immateriality.

- 95 627,23 (IV.x.17). The crucial words are 'pure', 'accidental', 'unguided', 'blind'. Locke's language is strongly reminiscent of Joseph Glanvill's argument from design, which expressly had reference to effects that he believed to be mechanical: 'For though *blind matter* might reach some elegancies of individual effects; yet specific conformities [i.e. species of animals and plants] can be no *unadvised* productions, but in greatest likelihood, are regulated by some *knowing* agent' (Glanvill 1661:45).
- 96 Hobbes himself claimed that he held only that, if we speak of God 'for honour', it must be as of something material, although his real nature is incomprehensible. Cf. Hobbes 1651:208 (*Lev.* III.34); 1839–45:IV 426ff.
- 97 Locke 1823:37. A similar passage on page 33 refers explicitly to *Essay* IV.x.16.
- 98 Cf. Locke 1975:297,26 (II.xxiii.5); 345,25 (II.xxvii.25).
- 99 541,8–15.
- 100 306,6 (II.xxiii.15). The passage is discussed in chapter 4, above.
- 101 Cf. 624,11–15 (IV.x.10).
- 102 See note 80, above.
- 103 Locke himself does not employ the terms 'nominal' and 'real' in relation to the general substances, matter and spirit. Cf. chapter 4, above.
- 104 Locke 1823:X 282ff.
- 105 Cudworth 1743:676. Cf. Locke 1975:716,23 (IV.xx.15).

15 *The Law of Nature and human freedom*

- 106 Locke 1975:229,17 (II.xx.2). For a possible, but not unequivocal qualification of this explanation of the difference between bodily and mental pleasures or pains, cf. 232,14 (II.xx.15).
- 107 351,19 (II.xxviii.5). My next paragraph summarizes succeeding sections.
- 108 353, footnote (first edition, II.xxviii.10): 'I call [it] the *philosophical Law*, not because Philosophers make it, but because they have most busied themselves to enquire after it.'
- 109 549,12 (IV.iii.18).
- 110 Hobbes offered perhaps the most obvious and famous example of a justification of moral rules in terms of their value to society, and so to the individual. But others, too, explicitly took the Golden Rule, the principle of social co-operation, as rational in virtue of self-interest. Hooker's justification of it along these lines is quoted, apparently with approval, at Locke 1963:310f, although Locke's own justification of it in the section following is rather different. Nevertheless Hobbes' role as a paradigm is reflected at Locke 1975:68,30 (I.iii.5), where Christian moral reasoning is compared both with the reasoning of a Hobbist and with that of the 'old Heathen Philosophers'.
- 111 Cf. 359,12 (II.xxviii.15).
- 112 King 1829:306ff. Colman 1983:171 finds the same general idea at Locke 1975:356,26, aptly commenting, 'The general picture is of a once universally acknowledged moral code, genuinely embodying the law of nature, which has...partially degenerated into the variable

law of opinion.’

- 113 Locke 1989:239 (sect. 185). Cf. Colman 1983:231ff.
- 114 Locke 1712: for Locke’s intention to publish, see Ayers 1980:102; Locke 1958:252ff.
- 115 Locke 1958:146ff. As von Leyden points out in his note, this view of reasoning was rejected at Locke 1975:595,2 (IV.vii.8).
- 116 Locke 1958:150ff; 204ff.
- 117 Although ‘reflection’ is treated with the senses in the *Essay* with respect to the acquisition of ideas, the corresponding ‘knowledge of existence’ is classified as ‘intuitive’ rather than ‘sensitive’. Cf. Locke 1975:618,20 (IV.ix.3).
- 118 Cf. 566,15 (IV.iv.8): ‘Nor are *Tally’s* Offices less true, because there is no Body in the World that...lives up to that pattern of a virtuous Man, which he has given us.’
- 119 Locke 1958:198f.
- 120 Locke 1975:516,26 (III.xi.16). Cf. Locke 1823:VII 112: ‘no one precept or rule is abrogated or repealed; nor indeed can be, whilst God is an holy, just and righteous God, and man a rational creature. The duties of that law, arising from the constitution of his very nature, are of eternal obligation.’
- 121 For a discussion of this and related points see Colman 1983:29–50 and 235–43.
- 122 Locke 1823:VII 4ff. He comments on the account of original sin as a punishment for Adam’s disobedience: ‘Could a worthy man be supposed to put such terms upon the obedience of his subjects? Much less can the righteous God be supposed, as a punishment of one sin, wherewith he is displeased, to put a man under the necessity of sinning continually, and so multiplying the provocation.’ The notion of a covenant was employed by some voluntarists to explain how an omnipotent God could be under any moral obligation at all. Boyle 1744:III 517f used it to explain how the Creator may in some cases oblige himself not to end the life he has given, but Boyle did not suggest that, without a covenant, God is above *all* rules of justice. Locke agreed that God could arbitrarily end life without injustice: ‘they owe it to his bounty; they could not claim it as their right, nor does he injure them when he takes it from them’. The implication of Locke’s argument is that this would hold even in the case of an extraordinary, perfectly good human being, and yet that it would be unjust for even ordinary virtue to be rewarded with eternal hell-fire.
- 123 Perhaps the nearest Locke got to allowing an absolute need for revelation is his statement in *The Reasonableness of Christianity* (Locke 1823:VII 139): ‘It should seem by the little that has hitherto been done in it, that it is too hard a task for unassisted Reason, to establish Morality in all its parts upon its true foundations; with a clear and convincing light.’ But in its rhetorical context this is by no means a categorical assertion, and the significance of the contrast with the greater optimism of the *Essay* (which Locke was still busily revising) should not be exaggerated. Cf. Colman 1983:138ff.
- 124 Locke 1958:180–9. Cf. Locke 1975:352,18 (II.xxviii.8). *Ius* and *potestas*, magisterial right and power, are distinguished but treated as necessary complements: *potestas* without *ius* is not really *potestas*, authority, but mere *potentia*, overpowering force. Similarly the right to govern presupposes the power to govern. Hobbes 1651:187f (*Lev.* II.xxxi) had expressly denied the importance of a *ius creationis*: ‘the Kingdom over men...belongeth naturally to God Almighty; not as Creator, and gracious; but as omnipotent. And though Punishment be due for Sinne onely,...yet the Right of Afflicting, is not always derived from mens Sinne, but from God’s power.’ This is the position Locke was rejecting.
- 125 Clarke 1706:69f. Quoted by Colman 1983:37.
- 126 It is easy to document this belief. Cf. Locke 1958:118ff (quoted by Colman 1983:30): ‘there

- is no fault, no guilt, where there is no law. Everything would have to depend on human will, and, since there would be nothing to demand dutiful action, it seems that men would not be bound to do anything but what utility and pleasure might recommend, or what a blind and lawless impulse might happen perchance to fasten on'. Or 'Ethica B', MS Locke c28, fol. 141: 'If man were independent he could have no law but his own will, no end but himself.'
- 127 Locke 1975:251, first edition reading (II.xxi.29). Cf. 257, first edition reading (II.xxi.33): 'that which at that time appears to him the greater Good absolutely determines his preference'.
- 128 253,12 (II.xxi.35).
- 129 254,27 (II.xxi.37f).
- 130 Cf. 281,25 (II.xxi.70).
- 131 275,20 (II.xxi.63).
- 132 267,18 and 270,4 (II.xxi.52 and 56).
- 133 268,13 (II.xxi.53).
- 134 281,17 and 268,2 (II.xxi.70 and 53).
- 135 282,32 and 283,12 (II.xxi.71).
- 136 279,7 (II.xxi.67).
- 137 244,24 (II.xxi.21).
- 138 267,5 (II.xxi.52). The previous quotation follows on directly.

16 Reflections on Locke's ethics

- 139 Cf. Dunn 1983.
- 140 Cf. Locke 1936:201f (journal entry, 15 July 1678) and Locke 1958:156f.
- 141 Bentham 1970:13.
- 142 Cf. Moore 1960.
- 143 Bentham 1970:11. Cf. Locke 1823:309 ('Of Ethick in General'): 'it being impossible to set any other motive or restraint to the actions of a free understanding agent, but the consideration of good or evil; that is, pleasure or pain that will follow from it'.
- 144 Nagel 1975.
- 145 Mill 1985:279–95.
- 146 It also seems possible to make sense of a sort of 'binding force' in relation to an ideal, e.g. of self-development, but the 'obligation' involved here seems more vulnerable to scepticism than our obligation in relation to others.

Part III: Identity (pp. 203–92)

17 Introduction to Part III

- 1 Locke 1975:542,8 (IV.iii.6).
- 2 416,5 (III.iii.13); 440,24 (III.vi.4). On shape, cf. 333,5 (II.xxvii.8).

18 Locke on 'masses of matter'

- 3 Cf. Aristotle 1985:523–7 (320a–322a). The interpretation offered here is deliberately conservative. Subordinate forms were a particular target of Boyle's, who cited the view that they can survive the extinction of the specific form of the whole (Boyle 1667:291–363).
- 4 54f (1979:45f).
- 5 319–24.
- 6 Locke 1975:328,3 (II.xxvii.1), 'Wherein Identity Consists'.
- 7 Cf. 200,15 (II.xv.8).
- 8 329,4 (II.xxvii.2), taking it that the determining relation referred to is spatio-temporal continuity.
- 9 For the difficulty, cf. 542,24 and 32 (IV.iii.6); for the necessity, 306,29 (II.xxiii.19ff).
- 10 307,25 (II.xxvii.2).
- 11 The doctrine was discussed by Suarez who attributed it to Henry of Ghent (cf. Thiel, forthcoming). It was also advanced by Ockham.
- 12 This passage (Locke 1975:330,14) is the only one in the *Essay* even to suggest that Locke was committed (as Boyle was not) to a strictly atomistic corpuscularianism. But the point can be read as merely a conceptual one: in so far as the minute particles retain unchanging superficies, so far are questions of identity about them unproblematic, since *ex hypothesi* where we have the same such particle we have the same substance or matter united as one thing. He is not, one supposes, arguing that any intelligible account of the identity of material things must presuppose strict atomism. That would be a very uncharacteristic *a priori* argument for a particular form of the 'corpuscularian hypothesis'. Rather, he is assuming strict atomism for the sake of exposition.
- 13 Taking the argument to 331,33 (II.xxvii.3–5).
- 14 Hobbes 1839–45:I 135–8 (*De Corp.* II.xi.7).
- 15 The much favoured notion of a 'stage' or 'phase' suffers from the same sort of incoherence as Locke's 'mass of matter'. It is supposed that a manstage, a temporal part of a man, can intelligibly be said to exist with the spatial boundaries and synchronic unity of a man, but for just the moment it takes to identify it (or be stimulated by it—cf. Quine 1965:28, *et passim*). But John-now or John-at-*t* is identical with John. Imposed temporal boundaries can limit a temporal part, not of a *man* (who, like every substance, exists all at once) but of that dependent, non-substantial entity, a man's *life-history*. Merely temporal boundaries are merely ideal boundaries. This point is not appreciated by those who imagine that the only theoretical constraints on ontology are those of formal logic. 'Phases' cause no merely formal problems, since they, and their purported relationship to things, can be dealt with by canonical part-whole logic, converting substances into events. See above, chapters 8 and 9.

19 Locke on living things

- 16 Compare Boyle's own comments on the *syncretis* and *diacrisis* of ancient atomism (see note 4, above). In fact Boyle's treatment of the identity-conditions of individuals is fairly off-hand, since he is primarily interested in chemical change and the identity-conditions of stuffs. Discussion of the 'generation, corruption and alteration' of individual objects is by

way of analogy.

- 17 Cf. Geach 1962. Such a model is embodied in Hobbes' statement, 'for it is one thing to ask concerning Socrates, whether he be the same man, and another to ask whether he be the same body', and in Locke's, 'so that truly they are not either of them the same Masses of Matter, though they be truly one of them the same oak, the other the same Horse' (Hobbes 1839–45:137; Locke 1975:330,27). Locke, however, immediately reverts to a different model: 'Identity is not applied to the same thing.' Perhaps the nearest he got to talking of relative perishing and coming to be is at 419,21 (III.iii.21), which is still some way off.
- 18 In effect, an argument brought against Geach in Wiggins 1967.
- 19 Locke 1975:331,8. Cf. 330,35 and 331,3 (II.xxvii.4).
- 20 These are the descriptive headings of, respectively, II.xxvii.4, 5 and 9.
- 21 Browne 1643:23. A 'Trinity' because three in one. Cf. Locke 1975:460,8 (III.vi.32).
- 22 Locke 1975:331,9 (II.xxvii.4).
- 23 335,16 (II.xxvii.9).
- 24 337,1 (II.xxvii.11).
- 25 336,25 (II.xxvii.10).
- 26 332,26 (II.xxvii.7). Cf. 332,14; 333,2; 335,3; 335,10 (II.xxvii.6, 8 and 9).
- 27 346,26 (II.xxvii.26). Cf. 516,28 (III.xi.16).
- 28 I here apply to this argument a phrase Locke used later, at 335,5 (II.xxvii.8).
- 29 Discussed at Boyle 1667:96 (1979:66). He remarks that the unity of the resultant whole can hardly be attributed to a substantial form or membership of a natural species, since graft and stock are commonly of different species. The point is worth a little reflection on the part of those who attribute an indispensable 'individuating' role to sortal concepts.
- 30 Mackie 1976:148:

The concept of identity is not one which can be made more or less stringent; identity is an all-or-nothing affair. Where there is room for relaxation and indeterminacy is in the individual concept to which identity is applied. If we join two lengths of wire by twisting the ends together, do we have a wire or do we still have two wires? We must decide what to say about this and we can decide either way. But what we say determines, with no nonsense, the correct answer to the question, 'Is this the same wire as that?' There cannot be a 'this wire' and a 'that wire' which are more or less the same.

If all that Mackie means is that any indeterminacy pertains to individuation, and that none attaches to the logical concept of identity, we can agree. Moreover, it is true that the kind of decision he says we 'must' make is sometimes called for *ad hoc*, depending on the point of some counting exercise, or in order to fix the significance of a count. But that does not imply that, when dealing with natural individuals, any indeterminacy in individuation is due to 'our concepts' and never to nature. For a different view, see Evans 1978. I take it that Evans's brief argument (criticized in Parsons 1987) is sophistical.

20 Forms of material unity

- 31 Wiggins 1980:205f. On living things, see especially 77–86; on artifacts, 86–99; for the thesis of sortal dependency, as cited, 60.
- 32 Cf. 86: ‘living substances [are] systems open to their surroundings but not in equilibrium with them’.
- 33 Cf. 25, footnote 12. Wiggins accepted in discussion that his explanation of the identity of substances over time by means of the notion of principles of activity and the like is as applicable to some events as to substantial things. He remarked that such notions were not intended to discriminate things from events, which is done by the notion of endurance, as explained in the present passage. That makes the title of his book, *Sameness and Substance*, a little mysterious, but in any case, as we have seen, the endurance peculiar to substances is intimately linked with other features of the category, not least materiality. See especially chapter 9, above.
- 34 The particles of water in a glass do not, of course, *just* lie together, but interact. Yet they do not cohere in the way required to form an individual substance—the way in which the particles of a block of ice cohere.
- 35 This does not prohibit cases like that of Theseus’ ship, discussed in chapters 18, above, and 21, below, in which the ship is reconstructed out of the same formally significant parts, not just the same matter or particles.
- 36 There is much more to be said about parts, e.g. on such questions as why it is so unnatural to think of ‘untails’ as distinct parts of cats, and so natural to think of tails that way. It seems odder to think of a living bullock (or even a dead one) as made up of actual *filets mignons*, T-bone steaks and the like than as composed of a heart, brain, shin-bones etc. Perhaps there are conditions under which Aristotle’s principle (i.e. that a detached hand is neither a ‘hand’ in the same sense as, nor ever identical with, an undetached hand) comes under so much pressure as to need some qualification, e.g. if the undetached part is as readily identifiable and/or separable and/or dispensable as a leaf on a tree. But it is noteworthy that this continuum is one of, as it were, *relative* material discreteness and unity—even in the living animal a bone has a relative material unity which marks it off from the surrounding flesh, and it is that, if anything, which makes it seem different from the notional uncut sphere. Single, unmar-shalled intuitions in this area cannot carry much weight, and objections to a broadly ‘Aristotelian’ approach seem not to come near to outweighing objections to the other two approaches identified here. (The present motive for adopting it is not, of course, at all the same as Aristotle’s.)

21 Artificial and other problematic objects

- 37 Locke 1975:331,23 (II.xxvii.5).
- 38 As at 463,18 (III.vi.39).
- 39 318,21 (II.xxiv.3).
- 40 The phraseology (‘artificial Draughts of the Mind...signified by one name’) is of course very like that of the accounts of mixed modes. And at 318,7 (II.xxiv.2) an analogy is explicitly drawn between ‘collective’ ideas and ideas of simple modes of number.

- 41 Very often, depending on the nature of the collection, one sense predominates, e.g. it is the more natural way of thinking to suppose that we have the same pack when, and only when, we have the same cards.
- 42 The argument is so familiar that there is little point in citations, but it is endorsed, e.g., in Perry 1970.
- 43 Cf. Shoemaker 1970:531.
- 44 Or at least it implies that, when we purport to suppose that a statue and its distinct component piece of clay (which *ex hypothesi* had different life-histories) might have had a shared life-history with all temporal parts in common (e.g. when we suppose that they might have been, not only formed, but also destroyed together), then we are really supposing that a statue and a piece of clay might have existed numerically distinct from *that* statue and/or *this* piece of clay (because identical with each other). Yet the claim that it is logically impossible that just *this* still existing piece of clay should have been smashed at the time it constituted *that* statue is surely paradoxical. Identity is made to hang on an accident *post facto*, which, although not contrary to the way we think about events, *is* deeply contrary to the way we think about substantial objects. We may not know until at least one of them is over whether this speech now being made and this welcoming ceremony now going on are identical—but then, what happens in the course of an event is not accidental to the event, as it commonly is to participating substances. Indeed, the principle that substances exist all at once means that questions of identity at any time (i.e. relating to individual substances existing at that time) can be settled with reference to that time only, without respect to other times, i.e. such questions are *prior* to questions whether coexisting substances have the same or different life-histories.

22 *Personal identity before the Essay*

- 45 Locke 1936:121ff (Journal, 20 February 1682). The references to Cudworth's *True Intellectual System* are on 118 (18 February 1682).
- 46 Descartes 1964–76:III.143 (to Mersenne, 6 August 1640); III.580 (to Huygens, 10 October 1642).
- 47 More 1662:188. Cf. Smith 1660:IV 82ff.
- 48 Tillotson 1722:II 128. Cf. Leibniz 1981:236.
- 49 Bodleian MS Locke f.7, p.107 (5 June 1683).
- 50 Cf. Boyle 1979:192ff ('Some Physico-Theological Considerations about the Possibility of the Resurrection').
- 51 Cf. Descartes 1964–76:163–9 (to Mesland 9 February 1645). Some versions of this view, Aristotelian and Neoplatonist, distinguished the animal soul or form responsible for the identity of the living body from the intellectual soul which is immortal in the full sense. Cf. Lee 1702:123; Felton 1725:13.
- 52 e.g. at 340,4 (II.xxvii.15) and 344,13 (II.xxvii.23).
- 53 Turner 1685:119. I am indebted to the admirable Thiel (1983) for drawing attention to this and other relevant writings.
- 54 128.
- 55 123.
- 56 154.

57 152.

58 Sherlock 1690:68f. Sherlock's book was published six months after the *Essay*, and contains a number of Lockean touches, e.g. the point that 'we know no more, what the substance of Matter, than what the substance of a Spirit is'. Yet he did not question the presumption of a metaphysical spiritual unity and continuity underlying and explaining the phenomenal unity of consciousness, although the latter is the only kind of spiritual unity we can clearly achieve. Very possibly the influence was mutual, since Sherlock's initial account of the various principles of unity was as follows: (i) in the case of an unorganized body, 'whether it be simple or compounded of different kinds of matter, that is One numerical Body, whose Parts hang all together' (cf. Locke 1975:330,14—II.xxvii.3); (ii) in 'organical Bodies' the principle is 'the Union of all Parts, which constitute such an organized Body' (cf. 330,35—II.xxvii.4); (iii) finally, the 'Self-unity of the Spirit', which has 'no Parts and no Extension neither, that we can know of, can be nothing else than 'self-consciousness' (cf. 335,13—II.xxvii.9). That Locke explained 'simple' bodies as atomic rather than homoeomerous and (more significantly) refused to assume an underlying metaphysical simplicity or partlessness of thinking things does not prove that he did not draw on Sherlock's order of exposition here as a model in preparing the new chapter for the second edition. It is at any rate very understandable that Stillingfleet should have associated Locke's theories with unorthodox views of the Trinity.

59 Sherlock 1690:83.

60 69.

61 Cf. Boyle 1979:194: 'Nor is it by the vulgar only that the notion of *identity* has been uneasy to be penetrated.' The idea that there is a general problem about the concept of identity seems to have been commonplace.

62 Locke 1975:109,26 (II.i.10).

63 110,19 (II.i.11).

64 110,34 (II.i.12).

65 516,26 (III.xi.16). Leibniz commented on this passage that a finite pure spirit would surely fall into the relevant class. 'Corporeal' would seem a slip, unless Locke had corporeality in mind as a condition of pleasure and pain.

23 Locke's theory of personal identity

66 Locke 1975:333,6 (II.xxvii.8).

67 332,9 (II.xxvii.6).

68 The belief in a lower form of immaterial soul as the principle of animal identity is alluded to in passing at 337,24 (II.xxvii.12), but as hardly deserving consideration.

69 A general analogy between life and consciousness was a commonplace embodied in the language of philosophy. So at Locke 1936:123 (Journal, 20 February 1682) it was said that matter and spirit 'may both lye dead and unactive, i.e. the one without thought, the other without motion'. Other, theological, overtones manifest themselves in a comment on Thessalonians II.v.23, without date. St Paul's famous distinction between spirit (*pneuma*) and soul (*psyche*) Locke took to be

not any distinction of substance, as if in some men there were two

substances whereof the one was the soule the other the spirit but different constitutions of the same (substance *deleted*) person the one as enjoying barely an animal life.... The other is a (principle *deleted*) state of a nobler...life derived from Jesus Christ by faith to believers where by they being united to him...are put into a state of immortality.

(Locke 1987:II 675)

This superior life of believers can hardly be consciousness, and it would be interesting to know whether Locke was here trying to bring to bear on the interpretation of St Paul the model developed in II.xxvii, or whether the chapter on identity evolved in part from thoughts about St Paul.

- 70 Locke 1975:341,14 (II.xxvii.17). Cf. 344,14 (II.xxvii.23): 'For whatever Substance there is, however framed, without consciousness, there is no Person.'
- 71 Cf. 345,28 (II.xxvii.25): 'But let Men according to their divers Hypotheses resolve of that [the substantial basis of consciousness] as they please.'
- 72 345,25 and 347,18 (II.xxvii.25 and 27).
- 73 335,29 (II.xxvii.10).
- 74 Hobbes 1839-45: I 137 (*De Corp.* II.xi.7).
- 75 Locke 1975:336,14. It is in general important to distinguish these two senses of 'thinking thing'. Here it is used, as the gloss makes explicit, for the underlying substance (whatever it may be). At 335,12 (II.xxvii.9), however, 'the same thinking thing' refers rather to the phenomenal self united by consciousness (cf. also 347,17—II.xxvii.27). The term 'substance', of course, suffers from the same systematic ambiguity as 'thing' in Locke's theory: an animal is a substance, but we can have the same animal without having the same substance (in the other sense of 'substance').
- 76 336,33 (II.xxvii.11).
- 77 Cf. Locke 1823:IV 324 (*Second Reply* to Stillingfleet): 'Thus it is no impropriety of speech to say, 'this body of mine, which was formerly strong and plump, is now weak and wasted; though in such a [strict and philosophical] sense as you are speaking here, it be not the same body'.
- 78 Cf. Locke 1975:337,28 (II.xxvii.13): 'Whether if the same thinking Substance (supposing immaterial Substances only to think) be changed, it can be the same Person.'
- 79 Bayle adopted a similar argument for a purely sceptical purpose:

the Conservation of Creatures is a continued Creation. How d'ye know but that God permitted this morning your Soul to relapse into nothing, which he had continued to create till then, ever since the first Moment of your Life? How d'ye know but that he has created another Soul modified as yours was (that is to say, with the Reminiscence which he would have reproduced, if he had continued to create the [original] Soul).

(Bayle 1710:IV 2621)

The introduction of the notion of conservation changes the issue a little, but it would be interesting to know if the argument came from Locke, or from some other (possibly common) source.

- 80 Locke 1975:338,13 (II.xxvii.13).
- 81 The seeming inconsistency has been made much of by several commentators, e.g. Mackie 1976:184; Flew 1968:163f.
- 82 Cf. Locke 1823:IV 305.
- 83 Cf. Thomas 1979:18ff. Thomas well criticizes common misunderstandings of Locke's argument, although her defence of it seems over-ingenious.
- 84 Locke 1975:341,18 (II.xxvii.17). In fact the distinction between being a part of a unified special *object* of consciousness and being a part of the *basis* of consciousness is neatly, if unconsciously glossed over in the discussion of the separated little finger. The finger is said, before separation, to be 'comprehended under' the consciousness of the self, and the case is later imagined of its being separated from the body with that consciousness 'going along' with it, when it would be 'the *Person*, the same *Person*'; i.e. the attached finger is considered as an *object* of consciousness, while the detached finger is imagined (abnormally) to be the *subject* or seat of consciousness. It seems possible that Locke was deliberately echoing Richard Overton's trenchant materialist joke:

For could there be a *Facultative Substance* (as that of the Soule is made) without its body; then...were his whole body quite burnt and consumed away, except his GREAT TOE; *he*, even *his Soule*, might as well live in his GREAT TOE, as before in his whole Masse; yea, better in that, then without all, as they childishly suppose.

(Overton 1643:15)

- 85 See, for example, Locke 1975:339,25; 340,13; 342,17 and 36; 344,10; 347,10 (II.xxvii.14, 15, 19, 20, 22 and 26). This is a point of agreement between Locke and at least some of his critics, e.g. Leibniz.
- 86 341,10 (II.xxvii.16).
- 87 The closeness of 'consciousness' and 'conscience' appears if we compare 344,10 (II.xxvii.22), '[each] shall receive his Doom, his Conscience accusing or excusing him', with e.g. 347,8 (II.xxvii.26), 'The Sentence shall be justified by the consciousness all Persons shall have, that they *themselves*... are the same, that committed those Actions.'
- 88 341,28 (II.xxvii.17). Cf. 'own' at 342,11 and 'by my consciousness make my own' at 345,21 (II.xxvii.18 and 24).
- 89 346,26 (II.xxvii.26). Cf. 341,11 (II.xxvii.16).
- 90 Cf. Behan 1979 and Matthews 1977:24–34.
- 91 Cf. Olivecrona 1974.
- 92 Locke 1963:328 (II.26).
- 93 Hobbes 1651:80 (*Lev.* I.xvi).

24 Contemporary reactions to Locke's theory

- 94 South 1693:71. Cf. Lee 1702:125, where the same point is glossed, 'Actions can't unite themselves; it must be the Agent that must do that: Just in that respect as 'tis in Bodies.'
- 95 Butler 1896:I 385.
- 96 Leibniz 1981:236. Butler made a similar appeal to our natural knowledge of the continuity of an identical subject (Butler 1896:I 396).

- 97 Perronet 1738:126: 'Mr. Locke surely knew full well, that *mad Men* were capable of imagining any thing whatever. But...he means only such a *real Consciousness* as renders Men *justly* liable either to Rewards or Punishments for their past or present behaviour.' Cf. Mackie 1976:184: 'Here we can surely say that the subject is merely imagining and not remembering the supposed actions.'
- 98 Locke 1975:342,23 (II.xxvii.20).
- 99 344,4 (II.xxvii.22).
- 100 721,5 (IV.xxi.4).
- 101 344,18 (II.xxvii.23).
- 102 Berkeley 1964:III 299 (*Alciphron* VII); Reid 1785:333ff. Berkeley's version of the objection not only came before Reid's, but is more complicated, since it was directed against a watered-down version of Locke's theory. 'In the same person,' *Alciphron* admits, 'some old ideas may be lost, and some new ones got; but a total change is inconsistent with identity of person'.
- 103 Locke 1975:340,4 (II.xxvii.15).
- 104 542,8 (IV.iii.6).
- 105 344,13 (II.xxvii.23).
- 106 Perronet 1738:7-9.
- 107 Law 1823:180.
- 108 184. The view that a person is a mode had been advanced earlier in Collins 1707.
- 109 An interesting word (Law 1923:185). Law, like many of his contemporaries saw the justification of punishment (even divine punishment) as utilitarian, but assumed that its utility is dependent on its power to express or communicate something to the recipient. Like speech, punishment has its due effect on the recipient through (in the sense of Grice 1957) its non-natural meaning.
- 110 Cf. Law 1823:186:

Let personality answer to a line or surface; let the substances it is predicated of, like the infinite variety of solids in nature, (with their appendages, heat, cold, colour, etc.), in which length and breadth are found, vary as you please; still the abstract ideas of line and surface, and therefore of person, will remain invariable.

Cf. Locke 1975:517,6 (III.xi.16).

- 111 Law 1823:187f.
- 112 199ff. The claim is of course itself quite wild, since 'person' is a noun predicable of a substance, as 'gratitude' is not. It is in this like (say) the noun 'noble', and 'personality' is like 'nobility'. Law himself used 'personality' frequently, if on the false supposition that 'same person' is equivalent to 'same personality'. For discussion of Locke's views on the relation between 'father' and 'paternity', see chapter 8, above.
- 113 Clendon 1710:17f.
- 114 Law 1823:197f, 192, 195.
- 115 Coward 1704:105 (cited in Berman 1985).

25 Neo-Lockean and anti-Lockean theories of personal identity in analytic philosophy

- 116 Cf. Descartes 1964–76:691–5 (to Elizabeth, 28 June 1643): ‘Everyone knows that he is a single person having both body and thought together, which are of such a nature that the thought can move the body and feel the things which happen to it.’ Everyone has experienced (*éprouvé*) the union in themselves.
- 117 Parfit 1985:119–306.
- 118 The point is that, even if memory is the criterion of personal identity, identity is preserved only if we can say that the arriving person genuinely remembers what the departing person had done, and we can only say that if we count the causality of Teletransportation as part of the functioning of a memory-mechanism (or faculty of memory). Parfit uses an analogy to argue that this is a ‘trivial’ step. Suppose that the normal mechanism of sight is functionally replaced in a blind animal by an artifice which appropriately stimulates the brain, giving rise to appropriate visual sensations. Despite the unorthodox causality, we would still think of the animal as *seeing* the things in front of it. This analogy is hardly strong enough, however, to bear the weight Parfit places on it. It is not as if *any* ‘unorthodox’ mechanism linking a state of affairs with appropriate visual sensations would justify our saying seriously that the state of affairs is seen. I have never seen Lenin, although I have seen a projected image of him. Similarly, I do not myself calculate all that my calculator deals with, or remember what I look up in my diary. More relevantly, if someone could telepathically hook on to my memories in such a way that it was subjectively as if he were remembering doing and experiencing what I had done and experienced, there would be a causal-cum-intentional link between my past experiences and his present ‘recollections’, but the faculty in question would be telepathy, not memory. In this case, as in Parfit’s story of the Brain-state Transference device, there is a spatial or spatio-temporal gap between distinct individuals with distinct faculties which no merely causal-cum-intentional link can fill. Cf. 208f.
- 119 275: ‘A person is like a nation.’
- 120 Clarke 1708:122.
- 121 Parfit 1985:297.
- 122 It ought to be admitted somewhere in this book that the present use of the word ‘abstract’ can cause confusion, since ‘abstract’ seems to have a degenerate sense for many people such that whatever is abstract is universal, or at least outside the natural world. In a perfectly good and traditional sense, however, such particulars as John’s sense of humour and the British constitution are ‘abstract’ as John and Britons are not.
- 123 An argument much like the present one has been independently arrived at by Paul Snowdon in an interesting recent paper (Snowdon 1990:91f).
- 124 Assuming the attractive principle that events with the same causes and effects (which are intersubstitutable in all causal conditionals *salva veritate*) are identical. Of course the notion of two thoughts or two speech acts (the man’s and the person’s) realized in but a single event might be felt sufficiently embarrassing for this principle to be called into question, but such a move promises nothing more advantageous than a disastrous complication in the philosophy of language and mind. Better, surely, to dump the distinction between man and person.
- 125 Locke 1975:342,28 (II.xxvii.20).

- 126 Parfit's clear recognition that on his theory a person is not a substantial object is not shared by all who advance similar views. Shoemaker's functionalist account, for example, is accompanied by the claim that it is 'trivially' true that a person is a substance in the same sense as a horse is a substance. That is only possible, however, because Shoemaker holds that the principle of unity and continuity of a horse is a matter of causal links no different in principle from the psychological links which bind together his 'persons', i.e. he ignores the materiality (etc.) of substances, and seems committed to holding that any such 'continuant' as a nation or a thunderstorm is a substance. His argument at least illustrates the point that an adequate theory of personal identity must rest on an adequate theory of substance (Shoemaker and Swinburne 1984:69–129, especially 72–6; cf. Shoemaker 1984, *passim*).
- 127 Cf. Parfit 1985:260. We could say of our case of multiple personality that 'Mary' and 'John' suffer from just this delusion; or, more perspicuously, that the man Richard Roe suffers from the continuing delusion that 'Mary' and 'John' are real people, i.e. are more than merely ideal or abstract objects realized in himself.
- 128 The present attempt to answer this question is, of course, itself theoretical to the extent that it is pursued by argument and the answer is set within a structured approach to epistemology and ontology. We cannot proceed simply by asking the reader to stop thinking and attend to the given. But it does involve an appeal to the reader to think reflectively about what experience is really like.
- 129 Kant 1963:342 (*Critique of Pure Reason*, A363).
- 130 373 (B415).
- 131 381 (B427f).
- 132 340 (A359f). I have mingled references to the first and second editions of the Critique, on the assumption that the two are compatible in this respect.
- 133 Shoemaker 1985:102f, cf. Wittgenstein 1964:66f. Of course it is true that when I remember doing something I do not enter into my memory in *just* the same way as other people do, e.g. I do not remember my looking a certain way, and I may remember how my arm ached. But that is far from meaning that I do not enter into my memory (or that I did not enter into the original experience) as a 'flesh and blood person', a material object.
- 134 Cf. Williams 1973:46–63.

Bibliography

- Aaron, R.I. 1963: *John Locke* (Oxford)
- Aarsleff, H. 1982: *From Locke to Saussure: Essays on the Study of Language and Intellectual History* (Minneapolis)
- Alexander, P. 1985: *Ideas, Qualities and Corpuscles* (Cambridge)
- Aristotle 1985: *The Complete Works of Aristotle: The Revised Oxford Translation*, ed. J.Barnes (Oxford)
- Arnould, A. and P.Nicole 1965: *La Logique ou l'Art de Penser*, ed. P.Claire and F.Girbal (Paris)
- Ayer, A.J. 1958: *Language, Truth and Logic* (London)
- Ayers, M.R. 1977: 'The Ideas of Power and Substance in Locke's Philosophy', in *Locke on Human Understanding*, ed. I.C.Tipton (Oxford)
- 1980: 'Locke's Translations of Nicole's *Essais*', *Locke Newsletter*, no. 11, pp. 101–3
- 1981: 'Mechanism, Superaddition and the Proof of God's Existence in Locke's *Essay*', *Philosophical Review*, vol. 40, no. 2, pp. 210–51
- Bacon, F. 1858–61: *Works*, ed. J.Spedding, R.L.Ellis and D.D.Heath (London)
- 1878: *Novum Organum*, ed. T.Fowler (Oxford)
- Bayle, P. 1710: *An Historical and Critical Dictionary* (London)
- Behan, D.P. 1979: 'Locke on Persons and Personal Identity', *Canadian Journal of Philosophy*, vol. 9, no. 1, pp. 53–75
- Bennett, J. 1987: 'Substratum', *History of Philosophy Quarterly*, vol. 4, no. 2, pp. 197–215
- Bentham, J. 1970: *An Introduction to the Principles of Morals and Legislation*, ed. J.H.Burns and H.L.A.Hart (London)
- Bentley, R. 1958: 'A Confutation of Atheism from the Origin and Frame of the World', in *Isaac Newton's Papers and Letters*, ed. I.B.Cohen (Cambridge)
- Berkeley, G. 1964: *The Works of George Berkeley*, ed. A.A.Luce and T.E.Jessop (London)
- Berman, D. 1985: 'Thinking Matter by J.W.Yolton' [review], *Philosophical Books*, vol. 26, no. 2, pp. 85–7
- Boyle, R. 1667: *The Origine of Formes and Qualities* (Oxford)
- 1685–6: *A Free Enquiry into the Notion of Nature* (London)
- 1744: *Works* (London)
- 1979: *Selected Philosophical Papers of Robert Boyle*, ed. M.A.Stewart (Manchester)
- Browne, T. 1643: *Religio Medici* (London) [Scolar Press facsimile]
- Butler, J. 1896: *The Works of Joseph Butler*, ed. W.E.Gladstone (Oxford)
- Clarke, S. 1708: *A Third Defence of the Immateriality and Natural Immortality of the*

Soul (London)

Clendon, J. 1710: *A Treatise on the Word Person* (London)

Coke, Z. 1654: *The Art of Logick* (London) [Scholar Press facsimile]

Collins, A. 1707: *Reflections on Mr Clark's Second Defence*

Colman, J. 1983: *John Locke's Moral Philosophy* (Edinburgh)

Copleston, F. 1962: *A History of Philosophy* (New York)

Coward, W. 1704: *Second Thoughts concerning the Human Soul* (London)

Cudworth, R. 1743: *The True Intellectual System* (London) [pagination as edition of 1678]

Davidson, D. 1984: *Truth and Interpretation* (Oxford)

Descartes, R. 1964–76: *Oeuvres de Descartes*, ed. C.Adam and P.Tannery (Paris)

—1970: *Descartes: Philosophical Letters*, ed. A.Kenny

—1985: *The Philosophical Writings of Descartes*, trans. J.Cottingham, R.Stoothoff and D.Murdoch (Cambridge)

Digby, K. 1645: *Two Treatises: Of Bodies and Of Man's Soule* (London)

Dummett, M. 1973: *Frege: Philosophy of Language* (London)

Duncan, W. 1748: *Elements of Logic* (London) [Scholar Press facsimile]

Dunn, J. 1969: *The Political Thought of John Locke* (Cambridge)

—1983: 'From Applied Theology to Social Analysis: The Break between John Locke and the Scottish Enlightenment', in *Wealth and Virtue*, ed. I.Hont and M.Ignatieff (Cambridge)

Evans, G. 1978: 'Can There be Vague Objects?', *Analysis*, vol. 38, no. 4, p. 208

Felton, H. 1725: *The Resurrection of the Same Numerical Body* (Oxford)

Flew, A. 1968: 'Locke and the Problem of Personal Identity', in *Locke and Berkeley*, ed. C.B.Martin and D.M.Armstrong (New York)

Ford, E.B. 1975: *Butterflies* (London)

Frege, G. 1968: *Foundations of Arithmetic*, trans. J.L.Austin (Oxford)

Gabbey, A. 1980: 'Force and Inertia in the Seventeenth Century: Descartes and Newton', in *Descartes: Philosophy, Mathematics and Physics*, ed. S.Gaukroger (Sussex)

Gassendi, P. 1658: *Opera Omnia* (Lyons)

—1981: *Institutio Logica 1658*, ed. H.Jones (Assen)

Geach, P. 1962: *Reference and Generality* (Ithaca, NY)

Gibson, J. 1968: *Locke's Theory of Knowledge and its Historical Relations* (Cambridge)

Glanvill, J. 1661: *The Vanity of Dogmatizing* (London)

Guerout, M. 1980: 'The Metaphysics and Physics of Force in Descartes', in *Descartes: Philosophy, Mathematics and Physics*, ed. S.Gaukroger (Sussex)

Harris, J. 1751: *Hermes* (London) [Scholar Press facsimile]

Hatfield, G. 1979: 'Force (God) in Descartes' Physics', *Studies in the History and Philosophy of Science*, vol. 10, no. 2, pp. 113–40

Hobbes, T. 1651: *Leviathan* (London) [Scholar Press facsimile]

—1839–45: *The English Works of Thomas Hobbes*, ed. W.Molesworth (London)

Hume, D. 1957: *Enquiries concerning the Human Understanding and the Principles of Morals*, ed. L.A.Selby-Bigge (Oxford)

—1968: *A Treatise of Human Nature*, ed. L.A.Selby-Bigge (Oxford)

- Jolley, N. 1984: *Leibniz and Locke* (Oxford)
- Kant, I. 1962: *Prolegomena to Any Future Metaphysics*, trans. P.G.Lucas (Manchester)
- 1963: *The Critique of Pure Reason*, trans. N.Kemp Smith (London)
- King, H.R. 1956: 'Aristotle without *Prima Materia*', *Journal of the History of Ideas*, vol. 17, no. 3, pp. 168–88
- King, P. 1829: *The Life of John Locke* (London)
- Kripke, S. 1972: 'Naming and Necessity', in *Semantics of Natural Language*, ed. D.Davidson and G.Harman (Dordrecht)
- Kuhn, T.S. 1962: *The Structure of Scientific Revolutions* (Chicago)
- Law, E. 1823: 'A Defence of Mr Locke's Opinion concerning Personal Identity'=J.Locke 1823: *The Works of John Locke*, III 177–201
- Lee, H. 1702: *Anti-Scepticism: or Notes upon Mr Locke's Essay* (London)
- Leibniz, G.W. 1875–90: *Philosophischen Schriften*, ed. C.I.Gerhardt (Berlin)
- 1981: *New Essays on Human Understanding*, trans. P.Remnant and J.Bennett (Cambridge) [pagination follows edition of A.Robinet and H.Schepers, Berlin 1962]
- Locke, J. 1712: *Discourses on the Being of a God*, [etc.] (London) [translation of part of P.Nicole, *Essais de Morale*]
- 1729: *Essai Philosophique concernant l'entendement humain*, trans. P.Coste (Amsterdam)
- 1823: *The Works of John Locke* (London)
- 1924: *An Essay concerning Human Understanding*, abr. and ed. A.S.Pringle-Pattison (Oxford)
- 1936: *An Early Draft of Locke's Essay*, ed. R.I.Aaron and J.Gibb (Oxford)
- 1958: *Essays on the Law of Nature*, ed. W.von Leyden (Oxford)
- 1963: *Two Treatises of Government*, ed. P.Laslett (London)
- 1975: *An Essay concerning Human Understanding*, ed. P.H.Nidditch (Oxford)
- 1976–88: *The Correspondence of John Locke*, ed. E.S.de Beer (Oxford)
- 1987: *A Paraphrase and Notes on the Epistles of St Paul*, ed. A.W.Wainwright (Oxford)
- 1989: *Some Thoughts concerning Education*, ed. J.W.Yolton and J.S.Yolton (Oxford)
- 1990: *Drafts for the Essay concerning Human Understanding and other Philosophical Writings*, Vol. I, ed. P.H.Nidditch and G.A.J.Rogers (Oxford)
- Long, A.A. and D.N.Sedley 1987: *The Hellenistic Philosophers* (Cambridge)
- Lovejoy, A.O. 1960: *The Great Chain of Being* (New York)
- McCann, E. 1983: 'Lockean Mechanism', in *Philosophy, its History and Historiography*, ed. A.J. Holland (Dordrecht)
- McCracken, C.J. 1983: *Malebranche and British Philosophy* (Oxford)
- Mackie, J.L. 1976: *Problems from Locke* (Oxford)
- Malebranche, N. 1923: *Dialogues on Metaphysics and on Religion*, trans. M.Ginsberg (London)
- 1980: *The Search after Truth*, trans. T.M.Lennon and P.J.Olscamp (Columbus, Ohio)
- Matthews, E. 1977: 'Descartes and Locke on the Concept of a Person', *Locke Newsletter*,

- no. 8, pp. 9–34
- Mellor, D.H. 1985: *Real Time* (Cambridge)
- Mill, J.S. 1985: *Utilitarianism*, ed. M. Warnock (London)
- Moore, G.E. 1960: *Principia Ethica* (Cambridge)
- More, H. 1662: *The Immortality of the Soule*, in *A Collection of Philosophical Writings* (London)
- 1743: *Divine Dialogues* (Glasgow)
- Nagel, T. 1975: *The Possibility of Altruism* (Oxford)
- 1979: *Mortal Questions* (Cambridge)
- Newton, I. 1662: *Philosophiae Naturalis Principia Mathematica*, trans. A. Motte (rev. F. Cajori) (Berkeley)
- Olivecrona, K. 1974: ‘Locke’s Theory of Appropriation’, *Philosophical Quarterly*, vol. 24, no. 96, pp. 220–34
- Overton, R. 1643: *Mans Mortalitie* (Amsterdam)
- Parfit, D. 1985: *Reasons and Persons* (Oxford)
- Parsons, T. 1987: ‘Entities without Identity’, in *Philosophical Perspectives 1: Metaphysics*, ed. I. Tomberlin (Atascadero, Cal.)
- Perronet, V. 1738: *A Second Vindication of Mr Locke* (London)
- Perry, J. 1970: ‘The Same F’, *Philosophical Review*, vol. 79, pp. 181–200
- Putnam, H. 1975: *Philosophical Papers II: Mind, Language and Reality* (Cambridge)
- 1981: *Philosophical Papers IV: Reason, Truth and History* (Cambridge)
- Quine, W.V.O. 1965: *Word and Object* (Cambridge, Mass.)
- 1969: *Ontological Relativity* (New York)
- Reid, T. 1785: *Essays on the Intellectual Powers of Man* (Edinburgh) [Scolar Press facsimile]
- Robinson, H.M. 1974: ‘Prime Matter in Aristotle’, *Phronesis*, vol. 19, no. 2, pp. 168–88
- Rogers, G.A. J. 1978: ‘Locke’s *Essay* and Newton’s *Principia*’, *Journal of the History of Ideas*, vol. 39, no. 2, pp. 217–32
- Russell, B. 1959: *The Problems of Philosophy* (London)
- Sherlock, W. 1690: *A Vindication of the Trinity* (London)
- Shoemaker, S. 1970: ‘Wiggins on Identity’, *Philosophical Review*, vol. 79, no. 4, pp. 529–44
- 1984: *Identity, Cause and Mind* (Cambridge)
- Shoemaker, S and R. Swinburne 1984: *Personal Identity* (Oxford)
- Sloan, P.R. 1972: ‘John Locke, John Ray and the Problem of the Natural System’, *Journal of the History of Biology*, vol. 5, no. 1, pp. 1–53
- Smith, J. 1660: *A Discourse Demonstrating the Immortality of the Soul*, in *Select Discourses* (London)
- Snowdon, P.F. 1990: ‘Persons, Animals, and Ourselves’, in *The Person and the Human Mind*, ed. C. Gill (Oxford)
- South, R. 1693: *Animadversions upon Dr. Sherlock’s Book* (London)
- Spencer, T. 1628: *The Art of Logic* (London) [Scolar Press facsimile]
- Strawson, G. 1989: *The Secret Connexion: Causation, Realism and Hume* (Oxford)
- Strawson, P.F. 1963: *Introduction to Logical Theory* (London)

- 1964: *Individuals* (London)
- 1966: *The Bounds of Sense* (London)
- Suarez, F. 1866: *Disputationes Metaphysicae* (Paris)
- Thiel, U. 1983: *Lockes Theorie der Personalen Identität* (Bonn)
- forthcoming: 'Individuation', in *The Cambridge History of Seventeenth Century Philosophy*, ed. M.R.Ayers and D.Garber (Cambridge)
- Thomas, J. 1979: 'On a Supposed Inconsistency in Locke's Account of Personal Identity', *Locke Newsletter*, no. 10, pp. 13–32
- Tillotson, J. 1722: *Works: Two Hundred Sermons and Discourses* (London)
- Turner, J. 1685: *A Discourse concerning the Messiah* (London)
- von Leyden, W. 1948: 'Locke and Nicole', *Sophia*, vol. 16, pp. 41–55
- Walters, S.M. 1961: 'The Shaping of Angiosperm Taxonomy', *New Phytologist*, no. 60, pp. 74–84
- Wiggins, D. 1967: *Identity and Spatio-temporal Continuity* (Oxford)
- 1980: *Sameness and Substance* (Oxford)
- Williams, B. 1973: *Problems of the Self* (Cambridge)
- Wilson, M.D. 1979: 'Superadded Properties: The Limits of Mechanism in Locke', *American Philosophical Quarterly*, vol. 16, no. 2, pp. 143–50
- Wittgenstein, L. 1958: *Philosophical Investigations*, ed. G.E.M.Anscombe and R.Rhees, trans. G.E.M.Anscombe (Oxford)
- 1964: *The Blue and Brown Books*, ed. R.Rhees (New York)
- Wolfson, H.A. 1977: *Studies in the History of Philosophy and Religion*, ed. I.Twersky and G.H. Williams (Cambridge, Mass.)
- Wright, J. 1983: *Hume's Sceptical Realism* (Manchester)
- Yolton, J.W. 1970: *Locke and the Compass of Human Understanding* (Cambridge)

Index

- Aarsleff, Hans 259n175
abstract and concrete terms 48f, 50–3, 231f, 254n80
abstraction, abstract ideas 3, 14, 45–7, 49, 55f, 58, 96, 276n110
abstract or notional objects 3, 15, 16f, 22f, 80, 82f, 86f, 91f, 96f, 102, 201–11, 230–2, 235f, 238f, 245, 249, 276n122, 277n127;
 see also distinctions of reason, modes accidents 21–3;
 real 28, 30, 49, 148, 171f;
 see also predicables, substance (and accidents)
action at a distance 129;
 see also gravity
actual properties, perspicuity of 56f, 127, 141f, 254n95;
 see also occult qualities, powers, mechanism
Adam 89, 113, 259n175
Alexander, Peter 252n43
altruism 167–9
ambiguity of artifact nouns 204–5, 208, 211f, 242;
 ascribed to ‘person’ 274–6, 289,
 see also persons, ‘I’;
 of collective nouns 241, 248;
 of geometrical nouns 243, 258f, 274, 289;
 transcategorical 241–5, 247–53;
 type-token 241–3;
 see also abstract objects, identity
ambiguous names 66, 77
analysis 25
angels 34, 124
animals status of 146f, 213, 252n49;
 thoughts of 9, 43, 254
a priori thought and knowledge 49f, 77, 80, 82, 88, 124, 131, 140, 159, 246
Aquinas, Thomas 112, 159, 160
archetypes *see* essences (in God’s mind), ideas (real and fantastical), substances (formation of ideas of)
Aristotle, Aristotelian philosophy 14–21, 24, 28, 30–2, 40, 49, 53f, 55–62, 64, 71, 72, 74, 80, 81, 89, 97f, 100f, 112, 123, 156, 158, 162, 173f, 175, 182, 184f, 187, 193, 201f, 218, 231, 250n14, 250n15
Arnauld, Antoine 48–9, 142
artifacts 60, 78f, 107f, 193, 201–12
atheism 144f, 150
atomism 1, 73, 107, 145, 269n12
atoms 56, 177

- Augustine 113f, 113
 axioms 157f
 Ayer, A.J. 5
- Beeckman, Isaac 116
 Behan, D.P. 275n90
 behaviourism 3, 6
 being chain or scale of 16, 57f, 144f, 147f, 153f, 255n101,
 see also perfection;
 idea of 55, 65, 303
 beings of reason *see* abstract objects, distinctions of reason
 Bennett, Jonathan 252n43
 Bentham, Jeremy 166
 Bentley, Richard 129
 Berkeley, George 8, 98, 132, 228, 229, 232, 238
 bodily sensations 239–40
 Boyle, Robert 14, 40f, 50, 55, 57, 73–4, 79, 118–9, 125, 128, 173f, 175, 178, 182f, 183f, 187, 243f,
 249n7, 255n118, 256n121, 267n122
 Butler, Joseph 168, 226f, 232
- Cambridge Platonists 37, 215;
 see also Cudworth
 categories 4, 7, 15, 16f, 44, 52–3, 55, 77 –
 changelings 59, 71f
 chemistry 55, 64, 74, 184
 chunks *see* identity (of simple concretions)
 Cicero 149, 231
 civil and philosophical language 63;
 see also language (reform of) civility 186
 civil law 155
 Clarke, Samuel 155f
 classifying and characterizing 95, 188;
 see also ‘what is it?’, species, taxonomy
 Clendon, John 231
 clocks, watches 22, 35f, 39, 59f, 79, 177, 187
 coexistence 8, 14, 31, 34, 82, 84, 122–30;
 see also necessary connection, substances (unity of)
 cohesion, coherence as a principle of individuation 73f, 107, 187–90, 199, 201, 240f, 270n34,
 see also identity (of masses of matter, of simple concretions), substances (unity of, as natural and
 ‘given’ individuals), unity (material);
 as a problem for mechanism 37, 148, 300f
 collections 99, 202f, 207f
 Collins, Anthony 153f, 276
 Colman, John 266f
 colour 81
 compound terms 20f, 48, 75f, 77–81, 99–102, 179–82, 204f, 211, 220, 229, 243, 246
 conceptualism 4–, 91, 93f, 95f, 99–103, 106f, 136f, 173, 184f, 186f, 193, 201, 202f, 211, 244f,

248fn2

- conceptual schemes 5–7, 53f, 95, 103, 137, 140, 248fn2
 ‘concretions’ *see* identity (of simple concretions)
 conscience 113, 223f, 228, 275n87
 consciousness 2f;
 as condition of immortality 254f;
 as principle of individuation 221, 255–85, 288–92, 322–6 *passim*,
 see also identity (personal);
 unity of 116f, 163, 205, 270, 286, 289
 constitution, real or internal 32–6, 58f, 61f, 64, 67, 70–3, 251n37;
 and artificial constitution 66;
 see also essence (real and nominal)
 continuants 89–91, 101–2, 193, 277n126
 Copleston, Frederick 261n3
 copula 255n87
 cosmological argument 144, 147f, 152f
 Coste, Pierre 51
 count-nouns 104, 194;
 see also substances (‘things’ and ‘stuffs’), identity (and names), individuation, nominalization
 creation 51, 112, 119, 143f;
 continuous 135, 139, 324
 criteria of identity and criteria of application 90f, 190f, 203, 260n176;
 stipulation of 107, **119–22**, **213–15**, 220f, **225–8**, 232, 245, 273f, **289**, **320n30**;
 see also individuation (principle of)
 Cudworth, Ralph 144–8, 153f, 213, 264fn81
 cupellation 63
 custom 27, 31f, 133f, 142
- Davidson, Donald 6f, 248n2
 death 74, 174–5, 180f, 182–3, 185, 188f, 213f
 definition real and nominal 17–21, 24f, 30f, 45f, 49f, 55, 57, 60, 70–1, 153f, 249n13,
 see also real and nominal essence, predicables;
 simple 22;
 stipulative 120f, 213–5,
 see also criteria of identity (stipulation of)
 deliberation 249–163
 Democritus 21
 density 23
 Descartes, René 14, 22–5, 27, 28, 32, 37f, 39f, 42f, 45–8, 53, 55, 97, 115–8, 128, 131, 134, 148,
 173, 184, 213f, 216, 218, 220f, 223, 231, 239f
 ‘desire’, meaning of 167f
 difference, *differentia* 18f, 55;
 see also predicables, definition
 Digby, Kenelm 23, 29–30
 distinctions of reason and real distinctions 3f, 7f, 14, 16f, 22f, 36, 45f, 55f, 95f;
 see also abstract or notional objects, modes
 division of linguistic labour 83

- dualism 34f, 37–40, 121f, 137, 144f, 152f, 214, 233f, 239–40, 253 *passim*, 260n180;
see also mind and body, matter (thinking), substance (immaterial)
- Dummett, Michael 248n2, 259n176
- Dunn, John 268n139
- duration *see* substances (endurance of), time
- egoism 167–8
- elect, the 112f
- emotivism 168f
- empiricism xvi, 4, 54, 93f, 155, 161, 235, 258n148;
 and rationalism **154–61**;
 ‘strong’ and ‘weak’ 154–6
- Epicurus 2, 183
- Escher 140
- essence, essences 17–20;
 in God’s mind 21, 132, 303,
see also universals;
 incorruptibility of 69f;
 and operation 43, 46, 167;
 real and nominal 38f, **40–2**, 52f, **56–8**, **66–76**, 78, 100–2, 144f, 226–8,
see also constitution;
see also substance
- eternal truths *see* universals, essences (in God’s mind)
- ethics 1f, 50, 77, 82, 112–4, 155–61, 165–9, 223;
 analogy with mathematics 187–91;
 and existence 188–90
- events 87–91, 99, 105, 210, 269n33;
see also continuants, modes, parts (temporal)
- evidence 2
- evolution 69–70
- existence dependent and independent 16–7, 22–4, 29, 94–6;
 in pieces 248f;
 potential 236;
 as *principium individuationis* 209f;
 supposed relative to classification 208, 216–19
- experience and theory 1, 93, 97f, 106f, 139, 245–7
- experiment 135f, 264
- extension alleged obscurity of 255n95;
 relation to matter 24–30, 52–5, 67;
 relation to substance 24–30
- Fall, the 112f
- Felton, Henry 272n51
- fictional beings 67f;
see also ideas (real and fantastical)
- Flew, Anthony 274n81
- Flood, the 129

- force 115–7, 124;
see also powers
- Ford, E.B. 70, 257n139
- form, specific or substantial as principle of individuation 17f, 173–5, 178f, 182, 193f, 196, 202, 207f, 219,
see also individuation;
 substantial 16, **19–21**, 25, 38, **65–71**, 94, 114–16, 147, 172, **206–8**, 304fn96;
see also species
- freewill 113f, 119, 162–3
- Frege, Gottlob 57, 99
- function, as principle of individuation 175, 201f, 206–8
- functionalism 3
- Gabbey, Alan 261n8
- Gassendi, Pierre 1f, 14, 24–5, 27–8, 32, 45
- Geach, Peter 182, 184, 234, 237, 269n18
- generation and corruption 175f, 182–4;
see also death, substantial change
- genes 70–2
- genus 18f, 55f, 69–70, 86, 106, 185;
see also species, predicables, definition, taxonomy
- geometry 19, 50;
 as analogy for ethics **57f**, 92, 134, **187–9**, 258f, 274;
 as analogy for physics 134, 144–6, 149;
 non-Euclidean 5
- Glanvill, Joseph 251fn36, 253n54, 265n95
- God ‘arbitrary will and good pleasure’ of 35, 112–4, 119f, 122–4, 127–30,
see also voluntarism, superaddition;
 as architect 21, 134, **137–9**, 144, 146, 148, 151f, **175–7**, **179–82**;
 as conserver of matter and motion 135–41, 146;
 death of 201;
 existence of 45–7, 169–83;
 as first cause 135, 171, 174;
 identity of 209, 222;
 immateriality of 44–6, 122, 174, 177–9, 315n94;
 as initiator of motion 173–5, 177;
 as source of obligation *see* ethics, voluntarism;
 whether subject to morality 132, 189f, 196f, 316n120, 317n122,
see also morality (binding force of)
- ‘good’ and ‘evil’, meaning of 155, 165f
- grace 113–4, 213
- grammar 4f, 44f, 51f;
see also nouns and adjectives
- gravity 120, 129f, 141, 144, 153, 261n25, 261fn35
- Grice, Paul 276n109
- Grotius, Hugo 157
- growth 19, 175, 196

Gueroult, Martial 117, 261n9

happiness 155, 161

Harris, Ian 260n2

Harris, James 104

Hatfield, Gary 261n9

hedonism 161, 166f, 168;

see also pleasure and pain

Heraclitus 175

Hobbes, Thomas 1, 24, 29f, 37, 40, 55, 97, 116, 125, 151, 155–8, 173, 178f, 182f, 185, 219, 224, 227, 267n124

Hooker, Richard 113, 157

Hume, David 5, 131–3f, 142f, 227, 229f, 235, 239, 241

'I', supposed referential ambiguity of 227f, 236–8, 240

idealism 53f, 53–140, 211, 235, 239;

see also conceptualism, pragmatism

ideas adequate and inadequate 35, 64, 87;

clear and distinct, obscure and confused **28**, 30, 32f, 37, 42, **47–50**, **75f**, **222**, 226;

collective 240;

real and fantastical 80f,

see also substances (formation of ideas of);

simple and complex 10, 31, 33–8, 57, 95, 97,

see also compound terms

identity 2, 5f, 8, 51f, 72, 100–2, 173–247, 248n2, 272n61:

of artifacts 127f, 229, 239–53;

of the body 207f, 212f, 218, 224, 256, 263,

see also the Resurrection;

criteria of *see* individuation (principle of);

and division 225, 233–6, 282f;

of 'finite intelligences' 209;

general theory of 208–10, 217–19;

and 'kinds' 208–10 216–20,

see also species (relation to individual);

of living things 207f, **210–33**, 235–7, 261f, 269f, 288f;

of 'masses of matter' **210–14**, 221f, 224, 234;

of modes and relations 120–2, 209f, 214, 262, 269, 273–7;

and names 208, 212, 222;

over possible worlds 104–6;

personal 47, 93f, 172, 189, **220–3**, **253–92**,

see also 'person', persons;

and place-occupancy 208–10, 217–19, 222, 224, 278, 309n180;

of representations 239, 242–4, 249–52;

of simple concretions 127f, 213, **229–38**, 239f;

of simple substances 210, 222;

of substance 221f, 261–3, 272;

supposed contingent 251, 321fn44;

- supposed relative 217–19, 236, 279, 284, 319n17;
see also individuation, unity
- imagination 25, 213
- imagism 23
- immortality 173f, 185, 213–4, 240, 253n59;
see also the Resurrection, identity (personal)
- impenetrability 51, 127f
- individuation 3–, 16–8, 20f, 58f, 73, 82–, 96f, 99–103, 173 –247*passim*;
 principle of 116, 205, 209–15, 228,
see also matter, form, existence, substance, cohesion, life, consciousness, soul (as principles of
 individuation), criteria of identity;
see also identity
- induction paradoxes of 133–4, 263n60;
 principle of 157–9, 161f, 313n61;
see also uniformity of nature
- intellect, pure 24, 25, 213, 239
- intentionality 3f, 213, 222
- internal constitution *see* constitution
- intuitionism 3
- ‘is’ of composition 210 –1
- ius creationis* 155, 157, 160, 165–7, 168, 267n124
- Jolley, Nicholas 257n134
- judgement 2
- Kant, Immanuel 4f, 7, 8, 54, 137–8, 159, 162, 168, 227, 235, 239 –40
- King, H.R. 250n15
- knowledge and belief 125;
 primary and secondary 2f;
 trifling and instructive 58;
 universal 143, 146f
- Kripke, Saul 66, 68, 71, 81, 85, 248, 257n134, 257n137
- Kuhn, Thomas 5
- language reform of 61, 63–4;
 and thought 3–11, 51–64
- Last Judgement 227, 235
- Law, Edmund 229 –31
- Law of Nature knowledge of 112, 114, 155–61;
 moral law 131–4, 153, 169, 287,
see also ethics;
 necessity of 189f,
see also voluntarism
- law of reputation 155–7, 168
- laws of nature and essences 168f, 168;
 and God’s essence 135–7, 139, 153f, 190;

- necessity of 146f, 163–7, 189f,
see also mechanism, necessary connection;
 physical laws 131f, 134f, 136–68 *passim*, 169,
see also voluntarism;
 and time and space 162f
- Lee, Henry 272n51, 274n94
- Leibniz, Gottfried Wilhelm 27f, 38, 42f, 54, 66, 89f, 98f, 123, 126, 142f, 226f, 246
- Leucippus 21
- Leyden, Wolfgang von 264n71
- 'liberty', definition of 264f;
see also freewill
- life as an accident 61, 173, 175, 188;
 as a perfection 173f;
 as a principle of individuation 206, 208, **210**, **219–25**, 231f, 258, 260–3, 265, 267, 269f, 288f;
 of the soul 172, 288f, 323fn69
- Locke, John *passim*;
 development of his thought 1f, 34f, 55, 186–92, 254–9, 303n77, 307n148, 314fn81, 315n88,
 322fn58
- logical form, of propositions 2–, 14, 51, 93f, 93
- Lovejoy, A.O. 255n101
- Lucretius 153
- lumps 107f, 179;
see also identity (of simple concretions)
- Luther, Martin 113
- McCann, Edwin 124–6, 128f, 262n36
- McCracken, C.J. 261n17
- Mackie, J.L. 234, 258, 274n81, 274n97
- Malebranche, Nicolas 15, 46–7, 54, 98, 118, 132
- mass-terms 103, 194;
see also substances ('things' and 'stuffs')
- materialism 32, 37–40, 144, 151, 152, 213–4, 228, 235;
see also matter (thinking), dualism
- material unity *see* unity
- materia prima* 21, 175, 250n14, 250n15;
see also matter (Aristotelian)
- mathematics 50, 77;
see also geometry
- matter Aristotelian **21f**, 23, 46f, 55, 96, 175, **250n14n15**, 253n75;
 passivity of 135, 139f, 155, 311;
 as principle of individuation **21**, **124**, 211–14, 221f, 224, **241**, **247f**, **250f**, 255f, 261f, 264,
see also individuation, paradoxes (of overlapping substances);
 as a substance **24–30**, **33–5**, 38, 40, **66f**, **116**, **125–7**, 181f, 294;
 thinking 142f, 150, 169–83, 261,
see also materialism
- Matthews, Eric 275n90
- meaning 5f, 276n109

- mechanism 21f, 40f, 49, 56–9, 78f, 98, 115–30, 131–2, 145, 175, 187, 232, 245f;
 definition of 135, 153;
 assessment of 115f, 125–7, 167f;
 forms of 135–53, 163f, 181f
- Mellor, D.H. 260n187
- memory 31, 81f;
 misremembering and forgetting 264–6, 270–4, 276, 325n97, 326n102;
 and personal identity 206, 254f, 262, 264, 266f, 279f, 287, 290, 327n118,
see also consciousness (as principle of individuation)
- metamorphosis 88, 193
- method 14, 135
- Mill, John Stuart 161, 168f
- mind and body interaction between 118, 119–24, 126;
 vital union of 257f, 262f, 276, 280, 287, 326fn116
- miracles 38, 114, 119f, 126, 129, 138, 145, 149
- modes 1, 23f, 45–50, 77–, 94, 115f;
 as abstract objects 120–2, 293,
see also abstract or notional objects;
 conventional unity of 98–103, 108f, 240,
see also unity;
 meaning of ‘mode’ 26f, 55, 91f, 111, 115;
 mixed 97–100, 321n40;
 natural and non-natural 91–3, 96–101, 308n158;
 simple *see* extension, geometry, number
- Molyneux, William 173
- monads 89f, 98
- monsters 74, 190
- Moore, G.E. 4, 166
- moral man, the 158f, 216f, 218, 222, 272n65
- morality binding force of 160f, 165–70, 267;
 secular *see* law of reputation;
see also ethics
- motion, transference of 51, 117, 124
- Nagel, Thomas 99, 168, 260
- ‘natural’, meaning of 21, 38, 143f, 149, 152, 199
- natural history 64, 139;
see also experience and theory
- naturalism, ethical, 166f
- natural kinds 20, 77–, 106f, 139f;
see also species, modes (natural and non-natural), essence (real and nominal)
- necessary connection 132, 121–30
- necessity, logical and natural 135f, 139–41, 264n66f
- Needham, Joseph 256
- Neoplatonism 216
- Newton, Isaac 51, 215, 262fn35
- Nicole, Pierre 157f, 264n71, 265n94

- Nietzsche, Friedrich 165
 nominalism 173f, 175, 179–80, 182, 184f, 187f;
 see also species, taxonomy
 nominalization 45f, 47–9, 52f, 95, 104, 231f, 275n112
 Norris, John 264
 nouns and adjectives 44–53, 77, 95, 188
 number 271n40
- obligation, and motive 161f, 163f;
 see also morality (binding force of) occasionalism 155
 occult qualities 55f, 123, 125, 255n95, 262fn35;
 see also powers, actual properties, mechanism
 Ockham, William of 113, 159f, 267n11
 Olivekrona, Karl 274n91
 omnipotence 113f, 118f
 ontological argument 148
 ontology 5, 7, 73, 93–6, 187, 203, 223, 229, 233, 236, 242f, 245, 246, 248f, 269n15
 original sin 269, 267fn122
 ostension, ostensive definition 4–6, 66–9, 71f
 ‘ought’, meaning of 155, 166
 Overton, Richard 243, 275n84
 ‘own’, ambiguity of 222
- pain 81;
 see also pleasure and pain
 paradox, paradoxes of attrition 196–7;
 of detached and undetached parts 235–7,
 see also parts;
 of induction 158, 313n60,
 see also induction;
 of overlapping substances **209–15, 218–24**, 227f, 234, 236, **239f**, 244, 247, **249–53, 282–4, 321fn44**,
 see also matter (as principle of individuation);
 of Theseus ship 211f, 239, 247f, 321n35,
 see also identity (of artifacts)
- Parfit, Derek 234–6, 238, 276fn118
 parts detached and undetached 17, 198–9, 271n36;
 temporal 6f, 102f, **105–8**, 121f, **124, 251f, 319n15**;
 and wholes 102f, 118, 124, 239, 247–51
 Pascal, Blaise 271
 passions 155, 156, 162–3, 168f
 Paul, St 272fn69
 Pelagius 113
 perceptual knowledge, authority of 1f
 perfections 126, 144–7, 152–4
 Perronet, Vincent 227, 229
 Perry, John 271

- 'person':
 meaning of 257f, 260f, 274f, 290;
 supposed, a forensic term 94, 205, 223, 266–8, 270–6;
see also persons
- personae* 224, 231, 236, 237
- personality, multiple 216, 227, 231, 233, 236–8, 243, 277n127;
see also persons, identity (personal)
- persons distinguished from men 185f, 216–8, 220, 226f, 229–38, 242–5, 276fn124,
see also moral man, the;
 supposed modes 273–7, 281f, 284f, 289, 295;
see also identity (personal)
- phase-sortals 87, 101f, 210;
see also compound terms, parts (temporal)
- phenomenalism 8
- philosophical explanation 5, 93f, 106f
- physicalism 99
- pieces *see* identity (of simple concretions), parts
- Plato 17, 40, 50, 160, 168, 218, 221
- pleasure and pain 266n106;
 as motives 191–5, 197–9;
 and value 184, 188f, 191, 197–9
- political theory 80, 155
- Porphry 18f, 55
- Port Royal *Logic* 47–9, 115, 125
- possibility, epistemic and ontological 14f, 126, 153, 219, 263n50
- powers, causal properties 14, 36, 64, 81f, 87–9, 105f, 116f, 119, 127, 141f, 263fn57;
see also occult qualities, actual properties
- pragmatism 137, 140, 233, 258f;
see also conceptualism
- precision, as an ideal 56f, 60, 74, 186, 190f, 269n30
- predicables 15f, 18–9, 49, 55–62, 256f;
see also genus, species, properties, accidents
- predication 3f, 14, 47f, 54f, 99
- primary and secondary qualities 8, 14, 34, 122f, 126, 145, 252fn60
- principium individuationis* *see* individuation (principle of)
- Pringle-Pattison, A.S. 252
- probability 134–6, 153
- properties 18–9, 35f, 57f, 59–62, 74, 88, 122, 143, 147, 150, 250n12;
see also predicables
- property rights, appropriation 166, 223–4
- psychological unity 233–8, 241–4;
see also consciousness (as principle of individuation)
- Pufendorf, Samuel 157
- Puritans 113
- Putnam, Hilary 66, 68, 71, 85, 248fn2, 257fn137
- quantities 74f, 103, 177–80

- quantity 21f, 23, 50;
see also matter, extension, number
- quantum mechanics 4, 106f, 136f
- Quine, Willard Van Orman 4–6, 248n2
- radical translation 5–6
- rationalism 114, 119f;
 relation to empiricism 154–6, 160f;
 ‘strong’ and ‘weak’ 154–6
- Ray, John 55
- real and nominal essence *see* essence
- realism 3–, 93, 95, 106, 132, 186, 193, 211, 248n2
- reason, rationality as characteristic of man 45, 61, 81, 216–8, 231, 254n79;
 and feeling 199–201
- reasons, as motives 161–3, 167–9
- reference 4f;
see also individuation, ostension
- reflection 27, 32, 218, 239f
- Reid, Thomas 228
- relations 49, 78f, 80, 230–1, 258;
 natural and non-natural 91–3, 96–101, 308;
see also modes
- Relativity, Theory of 4
- Resurrection, the 186, 214, 216, 218–21, 226–8, 235;
see also immortality, identity (of the body)
- revelation 1, 113f, 155, 228, 267n123
- reward and punishment and obligation 155f, 155f, 166f,
see also pleasure and pain;
 and personal identity 205, 264, 266, 270–4, 326n109
- ‘right’ and ‘wrong’, meaning of 155, 166
- Robinson, H.M. 251n15
- Russell, Bertrand 4f, 132–4
- scepticism ancient 2;
 about essences **28f**, **43–7**, 56, **63**, 71, 141, **147–50**, 152, 156f, 181, 205, **261**, 271, **324n79**;
 moral 196, 202
- Scholastics 28, 248n4, 250n7, 254n79f;
see also Aristotelian philosophy, Aquinas, Ockham, Suarez
- science 19, 30f, 49f, 77, 80, 127;
 and natural history 165;
see also definition, predicables, substances (as objects of science)
- second causes 115, 117, 119
- self-reference and self-awareness, the object of 236–40
- sensory field, unity of 238f
- Sherlock, William 215f, 226, 272fn58
- Shoemaker, Sydney 240f, 277n126
- signs, signification 3, 8, 41

- simple apprehension 28
- Sloan, P.R. 255n91
- Smith, John 272n47
- Snowdon, Paul 277n123
- Socrates 179
- solidity 51, 127f
- sortal terms, concepts 67, 75, 180, 182f, 189f, 193, 206, 270n29;
see also species, identity (and 'kinds')
- soul, spirit always thinks 37, 213f, 216;
 immateriality of 43–7, 116f, 122f, 142, 171–4, 177, 182, 205, 254f, 309n180,
see also dualism, materialism;
 meaning of 'spirit' 44f;
 as principle of individuation 255, 260–3, 276;
 as substance 29f, 38, 44f, 116f, 301n40, 302n54
- South, Robert 226
- space 4f, 7, 25, 41–3, 51
- species 17–9, 31–2, 40, 50, 52, 55–77, 86, 100f, 107f:
 relation to individual 20f, 66, 69f, 72–4, 86–9, 102, 127f, 206, 217–20, 224–8,
see also identity (and 'kinds');
see also predicables, taxonomy
- Spinoza, Benedict de 131, 173, 245
- spirit *see* soul
- Stillingfleet, Edward 27, 31, 36f, 38–9, 151–2
- Stoics 2
- Strawson, Galen 263n57
- Strawson, P.F. 7, 248, 263n61
- Suarez, Francisco 112, 157, 159, 250, 254n76, 267n11
- substance 14–, 245–7, et passim;
 and accidents 15–17, 26–30, 32–5, 37, 114f, 117, 137, 145, 299, 303n66;
 and essence 2, 15f, 20–3, 26–30, 34–6, 40f, 52–7, 66f, 173f, 181f,
see also essence (real and nominal);
 idea of substance in general 31f, 37f, 40f, 47–52, 55;
 immaterial *see* spirit;
 meaning of 'substance' 32, 34, 42, 48, 110–12;
 supposed unknowable 15–17, 28–30, 32–5, 37, 51–3
- substances activity of 97, 104–6;
 endurance of **104–8**, 121, **124**, 125, 319–22,
see also parts (temporal);
 formation of ideas of 31–8, 57f, 78–81, 98, 100f, 227;
 materiality of 4, 9, 101, 122–5, 251,
see also dualism;
 as natural or 'given' individuals 6, 11, **18f**, **25–7**, 63, **102f**, **107–9**, **112–14**, 117f, **122**, 124, **126–8**, **224–8**, **236–8**, 288f, 293, 295;
 as the objects of science **18–30**, 111f, 115f, **124–7**;
 simple and compound 210f, 261, 263–5, 269f;
 as subjects of predication **18–20**, 22–4, 28f, **54f**, **112–17**, 164f, 293;
 'things' and 'stuffs' 20, 61f, 80, 103, 111, 123;

- unity of 10, **33–7**, 40, 57, **87–9**, 98, **114–22**, 210–14, **220–8**, 231, 295,
see also unity, modes (conventional unity of)
- substantial change 16, 21, 58f, 174–5, 182, 186
- substrate, *substratum* 17, 24, 27–8, 30–2, 45, 90
- superaddition 38f, 122, 126, 143f, 146, 151–3
- suspension of desire 153–63
- synthetic *a priori* 137f
- taxonomy 63–4, 69–71, 86, 191, 203;
see also species
- teleology 19, 112, 117, 131, 148, 193
- teleological argument 148f, 153f, 157, 265n95;
see also God (as architect)
- telepathy 277
- theology and philosophy 112–4, 116f
- theoretical entities 106f, 140, 260n189;
see also experience and theory
- Theseus, ship of 178f, 201, 208f, 219, 270n35
- Thiel, Udo 267n11, 272n53
- things, in themselves 7, 53, 132
- ‘the thinking thing’, problematic
 reference of 262–4, 267, 275, 287;
see also persons, ‘I’
- Thomas, Janice 274n83
- time 4f, 6;
 and natural laws 162f
- transmigration of souls 216, 218, 221f
- transplant surgery 233f, 236f, 242f
- Trinity 114, 185, 215–6, 231f, 235, 273
- Turner, John 214–6
- uneasiness, as the only motive 161f
- uniformity of nature, principle of 118f, 132f, 135–8, 264n63
- unity material 107, 179–80, 186, 193–200, 240f;
 relation to continuity 210, 212f, 220–8, 260–3;
see also cohesion, substances (unity of), modes (conventional unity of)
- universals 4, 17f, 50, 52, 55–6, 58f, 112, 160;
see also nominalism, species, *a priori* thought and knowledge
- utility 158, 161
- voluntarism 36, 132f, 165f, 168;
 forms of 135–53, 163f;
 and theology 131–4;
see also Law of Nature, laws of nature
- Walters, S.M. 69–70, 257n138

'what is it?' 16, 20f, 24, 32, 53, 78f, 81, 205f

Wiggins, David 90, 105, 193–4, 248fn2, 260n183, 269n18, 269n31

will freedom of *see* freewill;

 weakness of 192–5

Williams, Bernard 278n134

Wilson, Margaret 123f

Wittgenstein, Ludwig 4f, 248n2, 278n133

Wright, John 263n57

Yolton, John 263n52