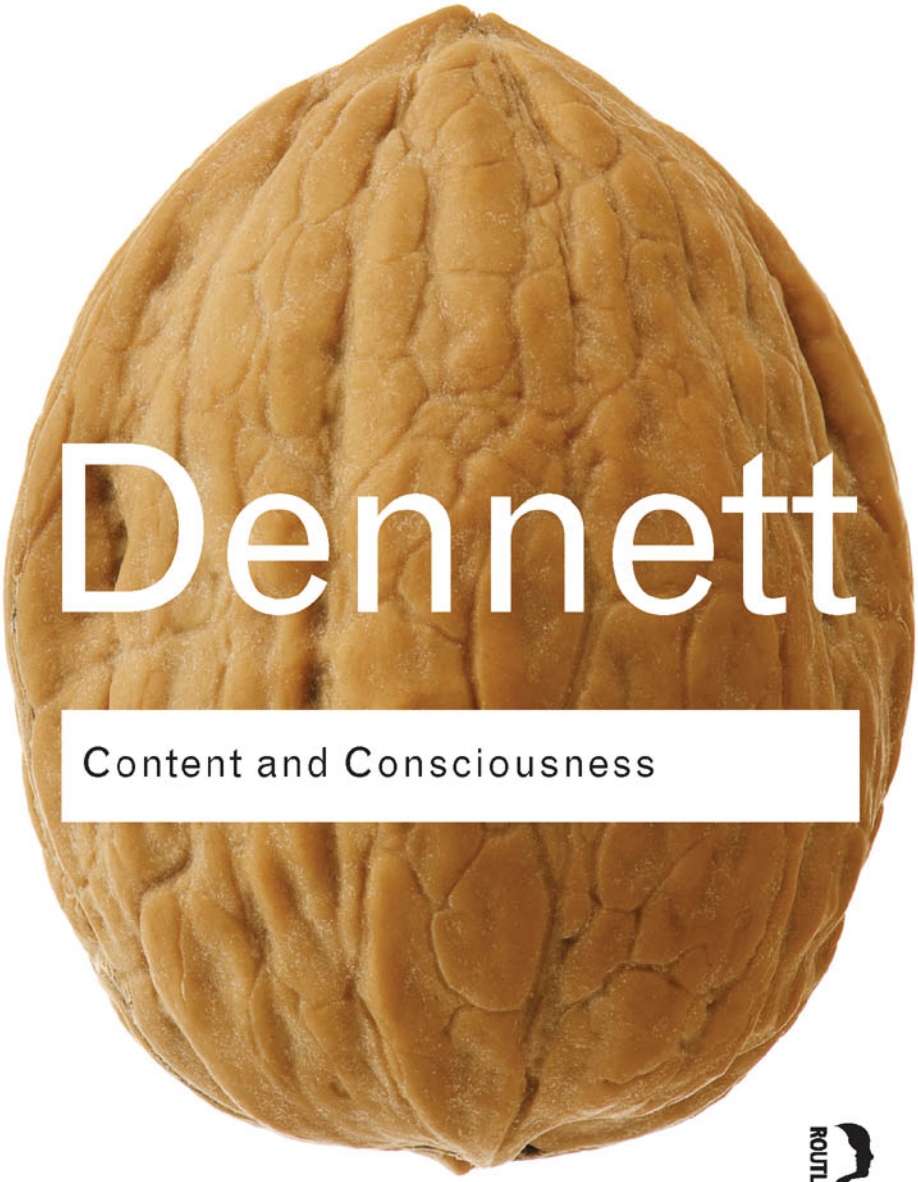


Dennett

Content and Consciousness

ROUTLEDGE

Content and Consciousness

'It presents a compelling and sometimes profound conception of the subject; it is ambitious without being grandiose; it employs scientific information effectively and ingeniously; its style, moreover, is enviably witty, clear, fluent and relaxed. One rarely encounters a difficult work of technical philosophy that is such a pleasure to read.'

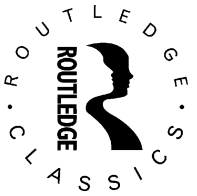
Thomas Nagel, Journal of Philosophy

'*Content and Consciousness* is an extraordinarily interesting and original book, and one which will raise the level of current discussion in the philosophy of mind.'

Richard Rorty, Philosophical Studies

'I have certainly been greatly stimulated by reading the book, and I recommend it to all others who have an interest in the problems of mind and body and of physicalism and its alternatives.'

J.C.C. Smart, Mind



Routledge Classics contains the very best of Routledge publishing over the past century or so, books that have, by popular consent, become established as classics in their field. Drawing on a fantastic heritage of innovative writing published by Routledge and its associated imprints, this series makes available in attractive, affordable form some of the most important works of modern times.

For a complete list of titles visit
www.routledge.com/classics

Daniel C.

Dennett

Content and Consciousness

With a new preface by the author

London and New York

First published 1969
by Routledge & Kegan Paul

Second edition published by Routledge 1986

First published in Routledge Classics 2010
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Avenue, New York, NY 10016

This edition published in the Taylor & Francis e-Library, 2010.

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 imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalog record for this book has been requested

ISBN 0-203-09295-3 Master e-book ISBN

ISBN 10: 0–415–56786–6 (pbk)
ISBN 10: 0–203–09295–3 (ebk)

ISBN 13: 978–0–415–56786–2 (pbk)
ISBN 13: 978–0–203–09295–8 (ebk)

For
Susan

Contents

Preface to the Routledge Classics edition

It has been a quarter century since I wrote the foreword to the first paperback edition. The cognitive science movement was well underway in 1984 and it is now going stronger than ever, with more and more philosophers playing the sorts of roles in it that I applauded then. Perhaps not so surprising is that in the interim something of a backlash has set in among a small coterie of philosophers who practise resolutely old-fashioned aprioristic philosophy of mind. I have yet to encounter anything worth more than a moment's reflection in that literature, but I grant that it is *logically* possible that something will emerge from it that other researchers on the mind will find worth their attention. I am not holding my breath.

Rereading my book, I have learned something curious about how my own thinking has changed. There are some passages that I find I no longer wish to endorse strongly, but at the same time don't wish to recant. I have moved from conviction to bland agnosticism, and find that I don't even have much to say about why my allegiance has waned. For this very reason, there

would be little or no point in my listing these passages; it would be of biographical interest at best and I don't wish to handicap them. So *caveat lector*; some of these passages may be best ignored or they may well be on to something important, may, for all I know, be gems I can no longer appreciate.

There is one set of ideas, however, mainly concentrated in Chapter 4, 'The Ascription of Content' that I do particularly want to commend to the reader. Here I set out for the first time (I think) the ideas that later became known as 'teleofunctionalism,' and drew out the implication that attributions of content were inextricably bound up with evaluations of aptness or appropriateness or rationality, an implication that later was dubbed and deplored as 'meaning holism' by Jerry Fodor. In spite of the fact that there has been a vigorous campaign by Fodor and others against these ideas, they continue to be reinvented. Recently, in *LOT 2: The Language of Thought Revisited* (Oxford University Press, 2008), Fodor has broadened his target and rechristened it *pragmatism*. If you think that 'abilities are prior to theories' or that '*knowing how* is the paradigm cognitive state and it is prior to *knowing that* in the order of intentional explanation' (2008, p. 10), then you are a pragmatist, according to Fodor's expansive definition, along with not just Dewey and Quine, but Wittgenstein, Ryle, Sellars, Putnam, Rorty, Dummett, Brandom, McDowell, Dreyfus, Vygotsky, Piaget, Bruner, and Gibson. Pretty good company, say I. All these folks have been utterly misled, Fodor insists, by their pragmatism, 'perhaps the worst idea that philosophy ever had' (p. 9). Doesn't that make you want to learn more? What sweeter encomium for my favourite idea could there be? Fittingly, the opposite of a pragmatist in this taxonomy is a Cartesian, a label Fodor now proudly sports. So here we have a fine confrontation. Which path should you take, the pragmatist or the Cartesian? I guess I should have belaboured my case for pragmatism more vigorously over the years, since there manifestly remain pockets of unpersuaded

philosophers. But the reason I didn't think I had to do this was that I thought I had sufficiently made the case in this book. I still think so.

Daniel Dennett
Cold Spring Harbor
June 1, 2009

Preface to second edition

It is now just twenty years since the first draft of this book was submitted (as a D. Phil. thesis at Oxford), and sixteen years since it was first published. In the intervening period the field of philosophy of mind has grown and changed enormously, a development that is perhaps made easier to see and appreciate by a reconsideration of the way the problems looked (to me) in the late 1960s.

When I was working on the book, its resolute naturalism and earnest concern with what science could tell us about the mind struck me as quite pioneering – or quite eccentric, depending on my mood. Philosophers of mind made something of a fetish of their distance from any empirical investigations, except of the most informal linguistic sort. Times have changed. Now we have cognitive science. There are now more than a few philosophers of mind who are vastly more knowledgeable about the brain than I was then (or am now). A fairly professional knowledge of the other cognitive sciences – psychology, artificial intelligence, linguistics – is now considered a virtual qualification for professional status in the discipline.

So what strikes me now about my book is not its pioneering stand, for we are almost all naturalists today, but its intermittent naiveté. This is mildly embarrassing, but not nearly so embarrassing as would be the discovery that I hadn't managed to achieve any advance in outlook over the years. There are also the unalloyed errors, of course, and these are indeed embarrassing. In fact, the only alterations to the text I have made, save for some typographical errors, are the elimination of a few preposterous howlers. (A good measure of what has been changed is the correction in example (4) in the second chapter: it was Ponce de Leon, not Hernando de Soto, who searched for the Fountain of Youth!) The more substantive errors, some of which I have still not recognized or recanted, alas, are left intact.

The task I set myself in the original Preface was 'to determine the constraints within which any satisfactory theory [of the mind] must evolve', and judging by the subsequent short span of theory evolution, I give myself high marks, at least for identifying the crucial issues and often even getting the matter right. For instance a voluminous debate on the identity theory has come – and gone – in the intervening years, leaving us with a residue of a few 'token-token' identities and a good deal of 'eliminative materialism', with the slack taken up by various sophisticated versions of 'supervenience' and accounts of psychology as an irreducible special science which is nonetheless properly deemed materialistic. In short, we are left with just about exactly the position I maintained (plus some useful sophistications) in Chapter 1.

Chapter 2, 'Intentionality', draws heavily on Chisholm and others, and introduces the problem of intentionality that has dominated much recent theorizing. There is little I would change in it today (except for those silly mistakes in the examples, which I have changed). The term 'intentional system' appears several times in the chapter and later in the book, but not with the precise sense I later developed for this term (in 'Intentional Systems',

Journal of Philosophy, 1971, and a number of other papers, all cited in the references of my *Elbow Room*, 1984). How fares 'centralism', the recommended theoretical approach to the problem of intentionality that consists in making an initial characterization of the phenomena to be studied in intentional terms, 'describing the events to be related in law-like ways using either ordinary, or semi-ordinary, or even entirely artificial Intentional expressions'? Except that no one calls his theory 'centralism', it fares well indeed, as the recent discussions of 'folk psychology' and its semi-ordinary and artificial alternatives in cognitive science attest. The debates on the ground rules have not diminished, with much attention being devoted to Putnam's and Fodor's methodological solipsism, its strengths, weaknesses, and rivals. This is one of the areas in which I have been provoked to embellish, adjust, revise and extend my thinking considerably – but not recant. In particular, my claims throughout the book about the relationships between inner and outer, function and meaning, rationality and meaning, and rationality and belief have been supported and wonderfully extended by a number of recent books, especially Ruth Millikan's *Language, Thought, and other Biological Categories* and Robert Stalnaker's *Inquiry* (both MIT Press, 1984).

Chapter 3, on 'Evolution in the Brain', stands up well, I think, in spite of its technical naiveté. The very recent upsurge in enthusiasm among neuroscientists for theories of learning as intracerebral evolution is particularly gratifying. While Edelman at Rockefeller, Changeux in Paris, the 'New Connectionists' in artificial intelligence and others are now developing 'evolutionary' models at a level of empirical detail and sophistication I could not imagine in the 1960s, I am pleased to see that their accounts appeal heavily to the concerns I outlined in this chapter. But it is also true that having said what I said, I simply didn't know what to do next with the ideas, so that the recent developments have opened new horizons for me.

The account of consciousness in Part II has some fairly dramatic

shortcomings, in my eyes. The account of introspective certainty has some important and salvageable points (in particular about the identity conditions of states and their relations to reports about them), but also some large confusions, which I have tried to correct in more recent writings. The distinction I draw in Chapter 6 between two different senses of awareness has been dropped from my later work on consciousness, not out of a conviction that it was entirely mistaken, but for strategic reasons: formulating it properly did not promise to be worth the time and effort. Recent discussions, however, have convinced me that *something like* that distinction is indeed a strategic necessity if various confusions are to be avoided, so I plan to refurbish a version of the awareness distinction in forthcoming work. Chapter 7, on mental imagery, has been almost entirely swept aside by subsequent empirical and theoretical work on the topic, but perhaps it is useful as an extremely simple and provocative introduction to the issues that are currently being explored. The last three chapters, on thinking and reasoning, action and intention, and language and understanding provide some foretastes of more recent discussions, and seem to me today to be not obviously wrong anywhere, but perhaps only because they are less detailed and ambitious than much current work on these central topics.

My own views of personal identity over time, and of responsibility, permit me to take a rather distanced and objective view of this book and its callow author. I find that all things considered I am glad it was written as it was when it was, and also glad that it is now being made available again, this time in a paperback edition. I learned quite a lot from rereading it, and hope that others will find it informative as well.

Daniel C. Dennett
Tufts University
March, 1985

Preface to first edition

Books attempting to tell the whole story of the mind have become rarer in recent years, for good reason. No one can hope to master the details, both of empirical data and of theoretical or conceptual nuance. In the face of staggering complexity, prudence has dictated to the student of mind that he must specialize – in the physiology of the nervous system, or in mathematical models of learning processes, or in the logic of key concepts such as belief, attention, or pain. This retreat from generality has been productive, but has left certain fundamental and pressing questions virtually untouched. What is the relation between a man's mental life and the events in his brain? How are our commonplace observations about thinking, believing, seeing, feeling pain to be mapped on to the discoveries of cybernetics or neurophysiology? These questions are important; their answers promise to bridge the specialties and consolidate their gains. But if attempts to answer them are confined, as they largely have been in the past, to philosophical guesswork on the one hand and the speculative perorations of retiring professors of neurology

on the other, no adequate answers will be forthcoming and there will be no unification of theory.

In examining these broad questions of mind and body, I do not try in this book to tell the whole story, but to set out the conceptual background against which the whole story must be told, to determine the constraints within which any satisfactory theory must evolve. The book specializes by slicing a cross-section, as it were, at a ninety-degree angle to the other specialties. Limiting the task in this way does not rule out all the risks of generality, however. Ideally, anyone hoping to work effectively in this area would have to keep abreast of half a dozen different scientific fields in addition to the advances in the philosophy of mind, but this is out of the question, so I have leaned heavily on résumés of research written for the non-specialist, scientists' gossip about current work, and especially the patient guidance of several colleagues in the different specialties. I have tried to couch all discussion of scientific matters in layman's terms – indeed it is only in layman's terms that I can understand it myself – and this course has side benefits as well as shortcomings. On the debit side, by the time any bit of science can be rendered in layman's terms it is usually a bit out of date, which, added to the time-lag of publication, isolates the discussion from the true frontiers of research. This has its bright side, however, for we should not want our working framework for theory to stand or fall on the often evanescent results hot off the presses of the learned journals.

Part I concentrates on the most general constraints governing scientific theories of the mind, and develops the notion of a distinct mode of discourse, the language of the mind, which we ordinarily use to describe and explain our mental experiences, and which can be related only indirectly to the mode of discourse in which science is formulated. In Chapter 1, a position of ontological neutrality is developed, which allows us temporarily to suspend decision of what ultimate ontological or

metaphysical shape our theory must take, materialistic, dualistic, interactionistic, vitalistic, etc. This allows certain sterile philosophical conundrums to be avoided, but leads directly to the most powerful challenge to unification in the theory of mind, the Intentionalist thesis that the mental mode of discourse is ultimately incompatible with the physical mode, and that no translations, reductions or unifications are logically possible. This challenge is examined in Chapter 2, and it is concluded that the best hope for unification lies with the development of a 'centralist' theory of mind. A centralist theory, in contrast to a peripheralist theory, would attempt to explain and predict human behaviour and experience by invoking central, internal states and conditions as crucial intervening variables in an explanation couched not in terms of mere stimulus and response, but in terms of purposive, conscious action. Chapter 3 determines some of the conditions of success for centralism and sketches a theory designed to meet these conditions. The essential task of centralism is seen to be justifying an interpretation of a physical system as a system whose states or events have meaning or content, and in Chapter 4 the conditions for such a justification are examined in detail, and the theory sketch is elaborated to meet these conditions. This leads to a general view of the relationship between the physical, mechanistic side of the story of the mind, and the non-mechanistic account embodied in our ordinary discourse about people.

In Part II, the bridge built in Part I is exploited in an analysis of consciousness, the feature of mind that is most resistant to absorption into the mechanistic picture of science. Chapter 5 gives a new account, based on the results of Part I, of the certainty of our introspective access to the 'arena of consciousness', and Chapter 6 analyses consciousness into several separable phenomena. Our ordinary view of consciousness is seen to be muddied by several sets of connotations that deserve separate treatment, and in Chapters 7, 8 and 9 these are given the

attention they deserve. Chapter 10 shows that certain unavoidable imprecisions in the formulation of centralism are inherent in that part of our given conceptual scheme that deals with people and their minds. There is a recurring theme running through the book that traditional analyses, both philosophic and scientific, have failed by postulating unanalysed elements having the very capacities to be analysed, thus postponing true analysis.

My considerable debts to a small number of writers will be evident from the frequency with which their names appear in footnotes. Others have helped more directly by reading drafts and making suggestions. First, my thanks and admiration go to Gilbert Ryle, under whose tolerant supervision the ideas for this book first took shape, and whose always insightful comments led me gently back from many false starts. Then to B. A. Farrell, Nicholas MacIntosh and J. Z. Young, who provided early guidance into the literatures of psychology and neurophysiology, and A. J. Ayer, Dennis Stampe and Jeffrey Sicha, who forced a number of my philosophical ideas into clarity. More recently, my colleagues and students at Irvine have provided valuable assistance, in particular, Gordon Brittan, Karel Lambert, James McGaugh, Julian Feldman, and Frank McGuinness. Ted Honderich's constructive criticisms of the penultimate draft led to many important revisions. I am indebted also to Henriette Underwood, Eva McCusker and Ida Brown for typing and excellent editorial suggestions beyond the call of duty. Part of my work on this book was supported by a grant from the University of California Humanities Institute, for which I wish to express my gratitude. I also wish to thank the editors of *Behavioral Science, Journal of Philosophy* and *Philosophy and Phenomenological Research* for permission to reprint with alterations parts of my articles published by them.

D.C.D.

Irvine, 1968

Part I

The Language of Mind

1

THE ONTOLOGICAL PROBLEM OF MIND

I THE MIND AND SCIENCE

Those who are convinced of the futility of philosophy are fond of pointing to its history and claiming that there is no progress to be discerned there. In no area of philosophy is this claim easier to support than in philosophy of mind, the history of which, when viewed through a wide-angle lens, appears to be a fruitless pendulum swing from Descartes' dualism to Hobbes' materialism, to Berkeley's idealism, and then back to dualism, idealism and materialism, with a few ingenious but implausible adjustments and changes of terminology. The innovations of one generation have been rescinded by the next so that despite a growing intricacy of argument and a burgeoning vocabulary of abstruse jargon, supplemented in each era by the fashionable scientific terms of the day, there have been no real and permanent gains.

The question that defined the pendulum is what the relation is

between mind and body, and the problem that set the pendulum in motion was Descartes' dilemma of interaction. If, as seems plausible at first glance, there are minds and mental events on the one hand and bodies and physical events on the other, then these two spheres either interact or not. The initially reasonable suggestion that they do interact leads, however, to an impasse of such difficulty that it can be held to be the *reductio ad absurdum* of dualism, at least of the Cartesian variety. If, *ex hypothesi*, mental events are non-physical, they can involve no physical energy or mass, and hence cannot in any way bring about changes in the physical world, unless we are to abandon the utterly central principle of conservation of energy and all its ramifications. Something must give way in this dilemma, and there are many choices available, all traced out by the swings of the pendulum. One can abandon the principle of conservation of energy, and this gives rise to the family of views of non-physical causes and 'occasions'; or one can preserve the principle and deny one of the other steps that lead to the dilemma. That is, one can deny that there are bodies and physical events and be an idealist, or deny that there are non-physical minds and mental events and be a materialist or physicalist, or hold for a dualism without interaction, and be a parallelist or epiphenomenalist.

The deficiencies of each of these alternatives, in each of their variations, have been well demonstrated time and again, but this failure of philosophers to find a satisfactory resting spot for the pendulum had few if any implications outside philosophy until recent years, when the developments in science, especially in biology and psychology, brought the philosophical question closer to scientific questions – or, more precisely, brought scientists closer to *needing* answers to the questions that had heretofore been the isolated and exclusive province of philosophy. Although one can still find in the current literature of the neurologists the old disclaimers about 'leaving to the philosophers' the 'mysteries' of consciousness, the 'initiation by the mind of neural