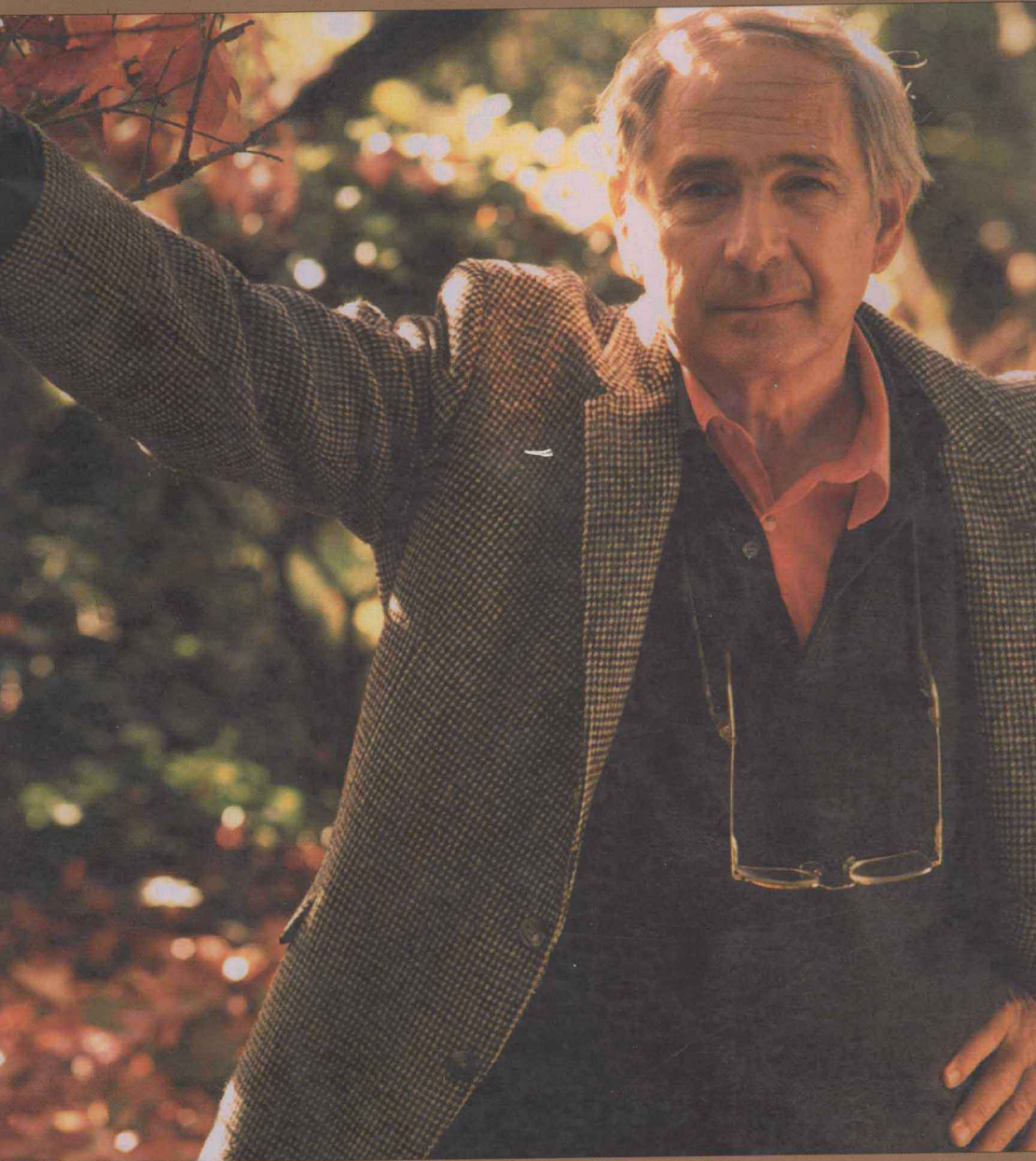


John R. Searle



CONSCIOUSNESS AND LANGUAGE

CONSCIOUSNESS AND LANGUAGE

JOHN R. SEARLE

University of California, Berkeley



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa
<http://www.cambridge.org>

© John R. Searle 2002

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2002

Printed in the United States of America

Typeface Baskerville 10/13 pt. System L^AT_EX 2_ε [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Searle, John R.

Consciousness and language / John R. Searle.

p. cm.

Includes bibliographical references and index.

ISBN 0-521-59237-2 – ISBN 0-521-59744-7 (pb.)

1. Consciousness. 2. Intentionality (Philosophy) 3. Language and
languages – Philosophy. I. Title.

B1649 .s263 c66 2002
126 – dc21

2001043890

ISBN 0 521 59237 2 hardback
ISBN 0 521 59744 7 paperback

For Grace

CONSCIOUSNESS AND LANGUAGE

One of the most important and influential philosophers of the last thirty years, John Searle has been concerned throughout his career with a single overarching question: How can we have a unified and theoretically satisfactory account of ourselves and of our relations to other people and to the natural world? In other words, how can we reconcile our common-sense conception of ourselves as conscious, free, mindful, rational agents in a world that we believe comprises brute, unconscious, mindless, meaningless, mute physical particles in fields of force? A cluster of individual questions that have preoccupied him – What is a speech act? What is intentionality? What is consciousness? What is rationality? – are all part of the larger problematic.

The essays in this collection are all related to the broad overarching issue that unites the diverse strands of Searle's work. The first five essays address the issue of how to situate consciousness in particular, and intentional phenomena in general, within a scientific conception of the world. The essays that follow discuss the implications of Searle's approach to the mind for psychology and the other social sciences, explore various ramifications of the theory of speech acts, and defend a version of mental realism by challenging the different forms of skepticism espoused by Quine and Kripke.

Gathering in an accessible manner essays otherwise available in relatively obscure books and journals, this collection will be of particular value to professionals and upper-level students in philosophy, as well as to Searle's more extended audience in such neighboring fields as psychology and linguistics.

John R. Searle is Mills Professor of Philosophy at the University of California, Berkeley. His previous Cambridge books are *Speech Acts* (1969), *Expression and Meaning* (1979), and *Intentionality* (1983).

CONTENTS

	Introduction	<i>page</i> 1
1	The Problem of Consciousness	7
2	How to Study Consciousness Scientifically	18
3	Consciousness	36
4	Animal Minds	61
5	Intentionality and Its Place in Nature	77
6	Collective Intentions and Actions	90
7	The Explanation of Cognition	106
8	Intentionalistic Explanations in the Social Sciences	130
9	Individual Intentionality and Social Phenomena in the Theory of Speech Acts	142
10	How Performatives Work	156
11	Conversation	180
12	Analytic Philosophy and Mental Phenomena	203
13	Indeterminacy, Empiricism, and the First Person	226
14	Skepticism About Rules and Intentionality	251
	<i>Name Index</i>	265
	<i>Subject Index</i>	267

INTRODUCTION

The essays collected in this volume were written over a period of two decades. They deal with a wide range of subjects and were intended for a variety of audiences. Despite the variety, there are certain unifying principles that underlie this collection; indeed, I have tried to make a selection that will exhibit a natural progression, as the topics move from consciousness to intentionality to society to language, and finally conclude with several debates about the issues that have preceded. In this introduction I want to try to state some of these unifying principles and offer a brief description (and note I say “description” rather than “summary” or “abstract”) of the essays.

There is a single overarching problem that has preoccupied me since I first began work in philosophy almost a half-century ago: How can we have a unified and theoretically satisfactory account of ourselves and of our relations to other people and to the natural world? How can we reconcile our common-sense conception of ourselves as conscious, free, mindful, speech-act performing, rational agents in a world that we believe consists entirely of brute, unconscious, mindless, meaningless, mute physical particles in fields of force? How, in short, can we make our conception of ourselves fully consistent and coherent with the account of the world that we have acquired from the natural sciences, especially physics, chemistry, and biology? The questions that have most preoccupied me – What is a speech act? What is consciousness? What is intentionality? What is society? What is rationality? – have all in one way or another been addressed to this larger problematic. I think this problem – or set of problems – is the most important problem in philosophy, and indeed there is a sense in which, in our particular epoch, it is the only major problem in philosophy.

If one accepts this characterization of my philosophical project then certain other features of my approach to these problems will become apparent, features that are by no means universally shared in the profession. First, philosophical problems of the kind we are dealing with should have clearly stated and definite solutions. There has to be a definite answer to such questions as "What exactly is consciousness and how does it fit in with the rest of the world?"; otherwise we are not making any progress with our task. Though philosophical problems have definite solutions, the solutions can seldom be given as direct answers to philosophical questions. The way I try to proceed is first to analyze the question. Indeed, this is the great lesson of twentieth-century linguistic philosophy: Do not take the questions for granted. Analyze the question before attempting to answer it. I like to proceed by analyzing the question to see whether it rests on a false presupposition, or whether it assimilates the problem at issue to an inappropriate set of paradigms, or whether the terms used in the question are systematically ambiguous. I find that in one way or another, philosophical problems characteristically require dismantling and reconstructing before they can be solved. Once clarity is achieved about exactly what questions are being asked, the answers, or at least the philosophical part of the answers, are often quite clear and simple.

Let me illustrate these points with some examples. Consider the famous "mind-body problem." What exactly is the relationship between consciousness and brain processes? To tackle this question we have to go behind the problem as posed and ask: What presuppositions does this formulation of the problem rest on? This problem resists solution as long as we continue to accept the traditional seventeenth-century vocabulary which presupposes that mental phenomena, naively construed, are in a completely different and separate ontological realm than physical phenomena, naively construed. Once we abandon this vocabulary, and the presuppositions on which it rests, the philosophical problem has a rather simple solution. Once we see that consciousness, with all its inner, qualitative, subjective, touchy-feely qualities, is just an ordinary property of the brain in the way that digestion is a property of the stomach, then the philosophical part of the problem is fairly easy to resolve. But there remains a terribly difficult scientific problem about how it actually works in the brain. I have something to say about both these issues, the philosophical and the neurobiological, in the course of these essays.

This is one pattern of addressing philosophical problems. A clarification of the problem will leave us with residual issues, but these are amenable to solution by scientific, mathematical, or other well-established means. What

I just said about the mind-body problem also applies, for instance, to the free will problem.

If one thinks of the primary task of philosophical analysis as giving a true, complete, and unifying account of the relations of humans and nature, then another distinction between this approach and that of much mainstream philosophy immediately appears. I find that I, personally, cannot take traditional skeptical worries very seriously. I think we made a mistake in taking so seriously, through the twentieth century, the line of skeptical argument and response that was begun by Descartes. I think in the seventeenth century it was reasonable to consider the existence of knowledge as problematic, and to feel that it required a secure foundation. Now it seems to me absurd to try to treat the existence of knowledge as problematic. If one thing is clear about the present intellectual situation it is that knowledge grows daily. The existence of knowledge is not in doubt. But you would be surprised at how much the persistence of skeptical worries, and the consequent epistemic stance, continue to have deleterious effects in philosophy. So, for example, Quine's famous indeterminacy argument, and indeed the whole project to examine meaning using the method of radical translation, is a matter of adopting an epistemic stance where I think it is inappropriate.

The first five essays – “The Problem of Consciousness,” “How to Study Consciousness Scientifically,” “Consciousness,” “Animal Minds,” and “Intentionality and Its Place in Nature” – are all concerned in one way or another with situating consciousness in particular, and intentional phenomena in general, within a scientific conception of how the world works. I think of consciousness and intentionality as biological phenomena on all fours with digestion or photosynthesis. If you think of mental phenomena in this way, you will not be tempted to think, for example, that a computational simulation of the mental phenomena is somehow itself a mental phenomenon. Computational simulations of the mind stand to the real mind the way computational simulations of the stomach stand to the real stomach. You can do a simulation of digestion, but your simulation doesn't thereby digest. You can do a simulation of thinking, but your simulation doesn't thereby think.

When I say that consciousness is a biological phenomenon, many people understand that claim to imply that you could not create consciousness artificially by building a machine. But that is exactly the opposite of the implication that I intend. The brain is a machine, and there is no reason in principle why we could not build an artificial brain that thinks and has other mental processes in the way that our brain does. Indeed, there is no reason in principle why we couldn't build an artificial brain out of completely nonbiological materials. If you can build an artificial heart that pumps blood,

why not build an artificial brain that is conscious? However, note that in order to actually produce consciousness, you would have to duplicate the relevant threshold causal powers that the brain has to do it. That is like saying that if you want to build an airplane that can fly, you don't necessarily have to have feathers and flapping wings, but you do have to duplicate the threshold causal powers that birds have to overcome the force of gravity in the earth's atmosphere. These are obvious points, but you would be surprised at how much difficulty I have had at getting them across to the philosophical and cognitive science community over the past two decades.

The first three essays are concerned explicitly with the problem of consciousness. They are arranged in chronological order and exhibit a certain level of development in my thinking about consciousness, together with increasing sophistication in neurobiology about the problem of consciousness. The first essay tries to state exactly what the problem of consciousness is by listing the main features that a theory of consciousness would need to explain. The second essay lists a number of mistakes that we need to avoid if we are to get a scientific account of consciousness, and the third essay reviews some features of the current project of explaining consciousness in neurobiology. In this last and most recent essay, I try to identify the prospects as well as the limitations of current research on consciousness. I distinguish between what I call the "building block model" and the "unified field" conception of consciousness, and make an argument that the unified field conception is more likely to succeed as a neurobiological research project.

This approach to the mind has important implications for the explanatory apparatus to be used in explaining cognitive phenomena. These issues are discussed in the next three essays. The first of these, "Animal Minds," defends the common sense view that animals have minds just as we do – with consciousness, beliefs, and desires, as well as pains and pleasures – but that their mental contents are restricted because animals do not have a language. "Collective Intentions and Actions" extends the account of Intentionality from the individual cases "I think," "I believe," "I intend" to the collective cases "We think," "We believe," "We intend." There is a nontrivial problem about getting the formulation of the conditions of satisfaction right for collective intentionality and much of that essay is devoted to that question. How exactly do we represent the content of my intention to do something, where I have the intention, as an individual, only as a part of our having the intention to do something as a collective? I am playing the violin part of Beethoven's Ninth Symphony and thus making my contribution to our collective activity of playing Beethoven's Ninth Symphony. You are, let us suppose, singing the soprano part and thus making your contribution to

our collective activity. How exactly do we represent the conditions of satisfaction of the individual and the collective intentions, and their relation to each other? I don't know if anybody else cares about that question, but it gave me a lot of headaches, and this essay is an attempt to characterize those relations in a way that recognizes both the irreducibility of collective intentionality and the fact that individual intentionality will be necessary to move my body even in cases where I move my body as part of a collective activity. I am definitely not trying to reduce the collective intentionality to individual intentionality, even though I recognize that all human intentionality is in the brains of human individuals.

The next two essays, 7 and 8, discuss the implications of my overall approach to the mind for psychology and other social sciences. In "The Explanation of Cognition," I explore in detail what I think is the correct explanatory apparatus for a sophisticated cognitive science to use, and in "Intentionalistic Explanations in the Social Sciences," I discuss the implications of my overall account of the mind for the traditional disputes between empiricist and interpretivist approaches to the problem of explanation in the social sciences.

The next two essays, 9 and 10, are concerned with extending my earlier work on speech acts in light of my researches in the philosophy of mind. A nagging dispute in the theory of speech acts is between those authors like Grice who take the individual intentionality of the speaker as the essential analytic device, and those like Austin and Wittgenstein who emphasize the role of convention and social practice. These appear to be inconsistent approaches, but I argue in "Individual Intentionality and Social Phenomena in the Theory of Speech Acts" that it is possible to reconcile these two approaches. Properly construed they are not competing answers to the same question, but noncompeting answers to two different and related questions. Essay number 10, "How Performatives Work," attempts to explain a phenomenon that originally gave rise to the whole subject of speech acts but which, oddly enough, seems to me not to have been satisfactorily explained. Namely, how is it possible that we can perform the act of promising or ordering, or christening, or blessing, and so forth, just by saying that we are doing it? The paradox here is that the whole subject of speech acts grew out of Austin's discovery of performative utterances and his ultimate rejection of the distinction between constatives and performatives. If you reject the distinction between constatives and performatives, as I do, and as Austin did – and without this rejection there is no such thing as a theory of speech acts – then you are still left with a problem. How do you explain the original existence of performative utterances? Essay 10 attempts to answer that question.

"Conversation" is about the possibility of extending my account of speech acts to larger stretches of discourse involving two or more people, to conversations. I reach a somewhat pessimistic conclusion. We will not get an account of conversation comparable in explanatory power to the theory of speech acts.

The last three essays are more argumentative than the earlier material. "Analytic Philosophy and Mental Phenomena" attempts to explain and overcome the puzzling tendency that many analytic philosophers have had to reject mental phenomena, naively construed. There has always been a tradition in analytic philosophy that attempts to get rid of consciousness and intentionality in favor of behaviorism, functionalism, computationalism, or some other version of "materialism." I diagnose and answer what I think are some of the most flagrant versions of this error.

Finally, the last two essays. In my work in both the theory of speech acts and the theory of intentionality I presuppose a form of mental realism. I assume we really do have beliefs and desires and other intentional states, and that we really do mean things by the words that we utter. Our mental states have a more or less definite content, and our utterances have a more or less definite meaning. When I began work on intentionality there were two forms of skepticism that seemed to challenge these assumptions: Quine's skepticism about the indeterminacy of translation and meaning, and Kripke's version of Wittgenstein's private language argument. If Quine's argument were correct, it would apply not only to linguistic meaning but to intentionality generally. Similarly with Kripke's skeptical interpretation of Wittgenstein's famous argument: If there is no definite ascertainable fact of the matter about whether or not I am using a word correctly, then it seems there is nothing for a theory of meaning and intentionality to be about. I could not proceed with my work on meaning and intentionality before I answered these two skeptical arguments, at least to my own satisfaction. I try to answer Quine's indeterminacy argument and to show that it is best construed as a *reductio ad absurdum* of the behaviorist premises from which it proceeds. In answering the Kripkean form of skepticism I distinguish two lines of argument in his book, and I claim that only the second of these is used by Wittgenstein. In any case I try to deal with both. This last chapter, by the way, has not been previously published because, frankly, I thought too much had been published about this issue already. I wrote it at the time of Kripke's publication but did not attempt to publish it. However, it does seem to fit neatly into the context of the present book, so I have included it, even though it is the only previously unpublished article in the collection.

THE PROBLEM OF CONSCIOUSNESS

I. What Is Consciousness?

Like most words, 'consciousness' does not admit of a definition in terms of genus and differentia or necessary and sufficient conditions. Nonetheless, it is important to say exactly what we are talking about because the phenomenon of consciousness that we are interested in needs to be distinguished from certain other phenomena such as attention, knowledge, and self-consciousness. By 'consciousness' I simply mean those subjective states of sentience or awareness that begin when one awakes in the morning from a dreamless sleep and continue throughout the day until one goes to sleep at night, or falls into a coma, or dies, or otherwise becomes, as one would say, 'unconscious'.

Above all, consciousness is a biological phenomenon. We should think of consciousness as part of our ordinary biological history, along with digestion, growth, mitosis and meiosis. However, though consciousness is a biological phenomenon, it has some important features that other biological phenomena do not have. The most important of these is what I have called its 'subjectivity'. There is a sense in which each person's consciousness is private to that person, a sense in which he is related to his pains, tickles, itches, thoughts and feelings in a way that is quite unlike the way that others are related to those pains, tickles, itches, thoughts and feelings. This phenomenon

This essay was originally published in *Experimental and Theoretical Studies of Consciousness*, CIBA Foundation Symposium 174 (Wiley, Chichester, 1993), pp. 61–80, copyright © 1993 John Wiley & Sons Limited. It is reprinted here with the kind permission of the CIBA Foundation and of John Wiley & Sons Limited. The theses advanced in this paper are presented in more detail and with more supporting argument in John Searle (1992).

can be described in various ways. It is sometimes described as that feature of consciousness by way of which there is something that it's like or something that it feels like to be in a certain conscious state. If somebody asks me what it feels like to give a lecture in front of a large audience, I can answer that question. But if somebody asks what it feels like to be a shingle or a stone, there is no answer to that question because shingles and stones are not conscious. The point is also put by saying that conscious states have a certain qualitative character; the states in question are sometimes described as 'qualia'.

In spite of its etymology, consciousness should not be confused with knowledge, it should not be confused with attention, and it should not be confused with self-consciousness. I will consider each of these confusions in turn.

Many states of consciousness have little or nothing to do with knowledge. Conscious states of undirected anxiety or nervousness, for example, have no essential connection with knowledge. Consciousness should not be confused with attention. Within one's field of consciousness there are certain elements that are at the focus of one's attention and certain others that are at the periphery of consciousness. It is important to emphasize this distinction because 'to be conscious of' is sometimes used to mean 'to pay attention to'. But the sense of consciousness that we are discussing here allows for the possibility that there are many things on the periphery of one's consciousness – for example, a slight headache I now feel or the feeling of the shirt collar against my neck – which are not at the centre of one's attention. I will have more to say about the distinction between the center and the periphery of consciousness in Section III.

Finally, consciousness should not be confused with self-consciousness. There are indeed certain types of animals, such as humans, that are capable of extremely complicated forms of self-referential consciousness which would normally be described as self-consciousness. For example, I think conscious feelings of shame require that the agent be conscious of himself or herself. But seeing an object or hearing a sound, for example, does not require self-consciousness. And it is not generally the case that all conscious states are also self-conscious.

II. What Are the Relations Between Consciousness and the Brain?

This question is the famous 'mind-body problem'. Though it has a long and sordid history in both philosophy and science, I think, in broad outline at

least, it has a rather simple solution. Here it is: Conscious states are caused by lower level neurobiological processes in the brain and are themselves higher level features of the brain. The key notions here are those of *cause* and *feature*. As far as we know anything about how the world works, variable rates of neuron firings in different neuronal architectures cause all the enormous variety of our conscious life. All the stimuli we receive from the external world are converted by the nervous system into one medium, namely, variable rates of neuron firings at synapses. And equally remarkably, these variable rates of neuron firings cause all of the colour and variety of our conscious life. The smell of the flower, the sound of the symphony, the thoughts of theorems in Euclidian geometry – all are caused by lower level biological processes in the brain; and as far as we know, the crucial functional elements are neurons and synapses.

Of course, like any causal hypothesis this one is tentative. It might turn out that we have overestimated the importance of the neuron and the synapse. Perhaps the functional unit is a column or a whole array of neurons, but the crucial point I am trying to make now is that we are looking for causal relationships. The first step in the solution of the mind-body problem is: brain processes *cause* conscious processes.

This leaves us with the question, what is the ontology, what is the form of existence, of these conscious processes? More pointedly, does the claim that there is a causal relation between brain and consciousness commit us to a dualism of 'physical' things and 'mental' things? The answer is a definite no. Brain processes cause consciousness but the consciousness they cause is not some extra substance or entity. It is just a higher level feature of the whole system. The two crucial relationships between consciousness and the brain, then, can be summarized as follows: lower level neuronal processes in the brain cause consciousness, and consciousness is simply a higher level feature of the system that is made up of the lower level neuronal elements.

There are many examples in nature where a higher level feature of a system is caused by lower level elements of that system, even though the feature is a feature of the system made up of those elements. Think of the liquidity of water or the transparency of glass or the solidity of a table, for example. Of course, like all analogies these analogies are imperfect and inadequate in various ways. But the important thing that I am trying to get across is this: there is no metaphysical obstacle, no logical obstacle, to claiming that the relationship between brain and consciousness is one of causation and at the same time claiming that consciousness is just a feature of the brain. Lower level elements of a system can cause higher level features of that system, even though those features are features of a system made up

of the lower level elements. Notice, for example, that just as one cannot reach into a glass of water and pick out a molecule and say 'This one is wet', so, one cannot point to a single synapse or neuron in the brain and say 'This one is thinking about my grandmother'. As far as we know anything about it, thoughts about grandmothers occur at a much higher level than that of the single neuron or synapse, just as liquidity occurs at a much higher level than that of single molecules.

Of all the theses that I am advancing in this article, this one arouses the most opposition. I am puzzled as to why there should be so much opposition, so I want to clarify a bit further what the issues are: First, I want to argue that we simply know as a matter of fact that brain processes cause conscious states. We don't know the details about how it works and it may well be a long time before we understand the details involved. Furthermore, it seems to me an understanding of how exactly brain processes cause conscious states may require a revolution in neurobiology. Given our present explanatory apparatus, it is not at all obvious how, within that apparatus, we can account for the causal character of the relation between neuron firings and conscious states. But, at present, from the fact that we do not know *how* it occurs, it does not follow that we do not know *that* it occurs. Many people who object to my solution (or dissolution) of the mind-body problem, object on the grounds that we have no idea how neurobiological processes could cause conscious phenomena. But that does not seem to me a conceptual or logical problem. That is an empirical/theoretical issue for the biological sciences. The problem is to figure out exactly how the system works to produce consciousness, and since we know that in fact it does produce consciousness, we have good reason to suppose that there are specific neurobiological mechanisms by way of which it works.

There are certain philosophical moods we sometimes get into when it seems absolutely astounding that consciousness could be produced by electrochemical processes, and it seems almost impossible that we would ever be able to explain it in neurobiological terms. Whenever we get in such moods, however, it is important to remind ourselves that similar mysteries have occurred before in science. A century ago it seemed extremely mysterious, puzzling, and to some people metaphysically impossible that life should be accounted for in terms of mechanical, biological, chemical processes. But now we know that we can give such an account, and the problem of how life arises from biochemistry has been solved to the point that we find it difficult to recover, difficult to understand, why it seemed such an impossibility at one time. Earlier still, electromagnetism seemed mysterious. On a Newtonian conception of the universe there seemed to be no place for the phenomenon of electromagnetism. But with the development of the theory

of electromagnetism, the metaphysical worry dissolved. I believe that we are having a similar problem about consciousness now. But once we recognize the fact that conscious states are caused by neurobiological processes, we automatically convert the issue into one for theoretical scientific investigation. We have removed it from the realm of philosophical or metaphysical impossibility.

III. Some Features of Consciousness

The next step in our discussion is to list some (not all) of the essential features of consciousness which an empirical theory of the brain should be able to explain.

Subjectivity

As I mentioned earlier, this is the most important feature. A theory of consciousness needs to explain how a set of neurobiological processes can cause a system to be in a subjective state of sentience or awareness. This phenomenon is unlike anything else in biology, and in a sense it is one of the most amazing features of nature. We resist accepting subjectivity as a ground floor, irreducible phenomenon of nature because, since the seventeenth century, we have come to believe that science must be objective. But this involves a pun on the notion of objectivity. We are confusing the *epistemic* objectivity of scientific investigation with the *ontological* objectivity of the typical subject matter in science in disciplines such as physics and chemistry. Since science aims at objectivity in the epistemic sense that we seek truths that are not dependent on the particular point of view of this or that investigator, it has been tempting to conclude that the reality investigated by science must be objective in the sense of existing independently of the experiences in the human individual. But this last feature, ontological objectivity, is not an essential trait of science. If science is supposed to give an account of how the world works and if subjective states of consciousness are part of the world, then we should seek an (epistemically) objective account of an (ontologically) subjective reality, the reality of subjective states of consciousness. What I am arguing here is that we can have an epistemically objective science of a domain that is ontologically subjective.

Unity

It is important to recognize that in non-pathological forms of consciousness we never just have, for example, a pain in the elbow, a feeling of warmth,