

A universe of zombies?

The problem of consciousness and the temptations of dualism

DAVID PAPINEAU

David J. Chalmers

THE CONSCIOUS MIND

In search of a fundamental theory
414pp. Oxford University Press. £22.50.
0 19 510553 2

The problem of consciousness has suddenly become one of the hottest items on the academic agenda. For most of this century it has scarcely been possible to mention the topic in polite academic society. But now consciousness is being widely billed as “the last frontier of science”, and thinkers from many different disciplines are competing to grab a piece of the theoretical action.

The challenge is to explain the relation between brain and mind. Science already tells us a great deal about the brain, about the way chemical transmitters enable banks of neurones to stimulate each other in increasingly well-charted ways. But science is silent about the lived world of conscious experience. It may tell us about neurones, but it does not tell us what it feels like to have a toothache, or to be angry, or to see something green.

A flurry of new ideas is currently on offer. Books defending different “theories of consciousness” are published each week. Journals of consciousness studies are starting up, the Internet is buzzing with discussion groups, international symposia convene regularly. *Time* magazine put the story on its cover earlier this year. True, it was the April 1 issue, but the eight-page piece “Can Machines Think?” was an admirably sober survey of current views of consciousness.

Many of the participants in the debate come from the hard sciences rather than psychology and philosophy. The topic seems particularly to attract eminent scientists hoping to cap their careers with a last big success. Francis Crick is only one of a number of Nobel Prize winners who are currently defending theories of consciousness, and they are followed by a legion of interested neuroscientists, molecular biologists and physicists.

Psychologists and philosophers are also thinking hard about consciousness. For many decades, these official theorists of mind have been combating various species of subjectivism, and as a result have had little sympathy for speculations about conscious experience. However, now that most of these battles have been won, even the psychologists and philosophers recognize that nagging questions about consciousness remain, and they too have plenty of ideas about how to answer them.

One of the most public figures in this debate has been David J. Chalmers, a young philosopher with a background in mathematics and cognitive science. His ideas about consciousness are often mentioned, and *The Conscious Mind*, his first book, has been widely heralded, with a preview in *Scientific American* last year, and prominent discussion in the *Time* piece. In the event, the result is disappointing. The book contains many sensible comments, but little is original, and the theory Chalmers himself defends is a variant of a familiar position with known deficiencies.

Chalmers is at his best on the failings of alternative “theories of consciousness”. Most current theories are reductive, equating consciousness with some objective feature of the brain. Hard scientists tend to go for physiological features; Crick, for example, identifies visual consciousness with 40 Hertz neuronal oscillations in the visual cortex. The psychologists and philosophers, by contrast, prefer more abstract structural features; their current favourite is the “informational role” played by internal sensory states. The obvious objection to all such theories, however, is that they quite fail to explain consciousness itself. As Chalmers points out, a reductive theory will at best identify the physical correlates of consciousness in the brain. But it won’t tell us why these brain states are associated with feelings. Why it should feel as it does – or indeed like anything at all – to have your visual cortex oscillate at 40 Hertz?

The point is not new, but Chalmers makes it well. He distinguishes the “psychological” and

solutions, to unravel the confusions that generate them. (We need “to show the fly the way out of the fly-bottle”.) This isn’t necessarily good advice for all philosophical problems, but therapy does seem a promising remedy for puzzlement about consciousness. Certainly Chalmers himself seems only to get further entangled when he attempts to tackle his “hard” problem straight on.

Chalmers’s solution depends on a dualist view of the mind. He argues that the conscious phenomenal properties of mental states are quite distinct from any physical properties. Dualism stands opposed to materialism, which holds that the phenomenal properties – the feelings – are not separate from the physical properties of the brain, but just the way the relevant brain states are experienced by the beings that have them. Chalmers borrows a nice image from the American philosopher Saul Kripke. Imagine that God creates all the physical features of the universe. Materialists will think that his work is

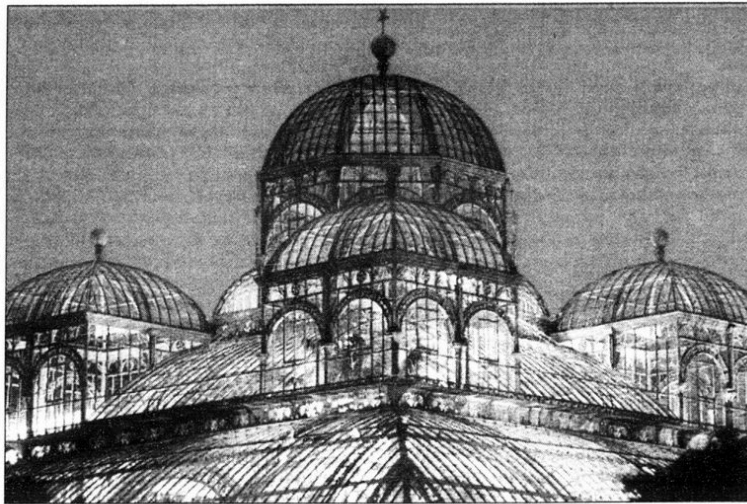
a zombie. So the fundamental issue between Chalmers and materialists is whether zombies are possible. Initial intuition certainly seems to favour Chalmers. There doesn’t seem to be anything incoherent in the idea of a zombie. Still, with so much at stake, we might well wonder whether initial intuition can provide adequate support for so weighty a metaphysical thesis as dualism.

This worry is amplified by the details of Chalmers’s position. On the surface, his dualism seems perfectly sensible. He is no new-age mystic. He wants to develop an objective science of phenomenal properties, and draws an analogy with the nineteenth-century recognition of electromagnetic forces as fundamental ingredients of nature. In line with this, much of the book is concerned with possible laws governing the emergence of phenomenal properties, akin to Maxwell’s laws governing the behaviour of electromagnetic fields.

But underneath, things are not so simple. The central problem facing any contemporary dualist is that twentieth-century science denies any causal powers to unreduced phenomenal properties. Phenomenal properties differ in this respect from electromagnetic forces. Electromagnetic fields have physical effects. They alter the motion of certain kinds of material particles. But it would fly in the face of modern physical science to suppose that Chalmers’s irreducible phenomenal properties have such causal powers. This would in effect postulate an extra mental force alongside the fundamental physical forces of gravity, the electromagnetic force, and the strong nuclear force. This might once have made sense, but the cumulative evidence of two centuries of physiological research weighs heavily against it. Certainly most physicists would be flabbergasted if it turned out that material particles – molecules in our brains, maybe – are sometimes accelerated, not because of any known physical forces, but because of irreducibly mental causes.

Some philosophers are able to turn a blind eye to this argument, but Chalmers is enough of a scientist to recognize its strength. So he opts for a form of “epiphenomenalism”, according to which conscious events are causal “danglers”, produced by brain processes, but not themselves able to exert any influence on the physical world. This is not an attractive position. It would have us deny that pain causes us to pull our hands out of fires, or that anger makes us red in the face. But Chalmers is prepared to bite this bullet. Indeed, it is to his credit that he positively emphasizes the counterintuitive consequences of his epiphenomenalism. For example, he accepts that a zombie duplicate of David Chalmers would behave physically just like the real Chalmers, since its absence of feelings would make no difference to the movements of its body. He points out some surprising implications.

[This zombie] talks about conscious experience all the time – in fact he seems obsessed by it. He spends ridiculous amounts of time hunched over a computer, writing chapter after chapter on the mysteries of consciousness. He often comments on the pleasure he gets from certain sensory qualia, professing a particular love for deep greens and purples. He frequently gets into arguments with zombie materialists, arguing that their position cannot do justice to the realities of conscious experience. And yet he has no



The Congo House, Laeken, Brussels, at night; from *The Glasshouse* by John Hix
(239pp. Phaidon. £45. 0 7148 3211 1)

“phenomenal” aspects of mental states. The “psychological” aspect of a mental state is its causal role in our cognitive workings, the way it interacts with other states in directing behaviour. The “phenomenal” aspect is the conscious feeling associated with the state, the what-it-is-like that we experience directly in our own case. Using this distinction, Chalmers contrasts two different problems of consciousness. The “easy” problem is to account for *psychological* properties in physical terms. Reductive theories of consciousness aim to do this by showing us how the relevant psychological roles are realized in our brains. But the “hard” problem remains. This is the problem of explaining why there should also be *phenomenal* properties associated with psychological ones, and here existing theories offer no help.

Chalmers is undoubtedly right to observe that reductive theories do not address his “hard” problem of consciousness. Whether they ought to is another question. Consciousness is a tricky business, and it is not necessarily wise to take its problems at face value. Wittgenstein held that philosophical problems need therapy rather than

done. But dualists will think he still has more work to do, to add the feelings to the physical facts.

Chalmers argues at length in favour of dualism. Much of his argument is technical, and I shall come back to some of the technicalities below. But the central idea is simple enough. As the creation story shows, we seem able to make good intuitive sense of a physical universe where the feelings have been left out – a universe of “zombies”, as Chalmers puts it, inhabited by beings who are physically identical to ourselves, but who have no conscious feelings. But if this zombie universe is possible, then this in itself shows that conscious phenomenal properties must be distinct from any physical properties. After all, it is precisely the absence of any extra phenomenal properties that would make the zombie world different from ours.

This is not a silly argument, but it can be rebutted. Materialists will simply deny that the zombie world is possible. They will say that once God has fixed the physical facts, he has fixed the phenomenal facts too. Even God could not make

conscious experience at all! Chalmers makes a brave attempt to make sense of this curious doppelgänger, but by this stage some readers may be wondering whether we ought not to have another look at materialism. At first blush, dualism may seem preferable to materialism. But if the cost is the causal impotence of consciousness, then materialism may be the better option after all. For one thing, materialism will restore the causal efficacy of conscious feelings, since it identifies feelings with their physical correlates. Your pain, according to the materialist, is identical with one of your brain states, and so will have the same physical effects as that brain state. Similarly, materialism has no problem explaining why Chalmers's doppelgänger is so interested in conscious experience. For it implies that any physical duplicate of Chalmers will automatically have his conscious experiences too.

Chalmers's initial argument against materialism started from the intuition that zombies are possible, and that conscious feelings are therefore distinct from any physical properties. This is indeed a strong intuition, but the difficulties it gives rise to should make us suspicious. Perhaps this intuition is a kind of illusion. Maybe we are seeing things from within a fly-bottle, which makes it look as if conscious and physical properties are distinct, when in fact they are not.

Here is what I think is going on. Mental states are peculiar, in that we can think about them in a special way. We can simulate them internally, in addition to thinking about them externally. Most things can be thought about only externally, since most things are external to our minds. But with pains, emotions and other mental states, we have an extra option. We can activate that part of the brain which is involved in feeling pain, and thereby think about the pain by literally recreating it.

There is plenty of physiological evidence to show that this is indeed how we imagine pains and other conscious states. Modern brain-scanning techniques show that we simply switch on the relevant parts of the brain. Of course, we don't switch on quite the same parts as are involved in actual pains – imagined pains don't feel as bad as real pains. But there is a large overlap in brain activity, which makes sense of the fact that imagined pains to a slight extent duplicate the unpleasantness of real pains.

Of course, we are also able to think about mental states, like anything else, from an external point of view. For example, we can imagine that the pain centre in someone's brain is being

stimulated. This means that we are able to think about mental states in two quite distinct ways – simulationally and observationally. And this then suggests a natural materialist explanation of why dualism seems intuitively true, even if it is false. The two ways of imagining mental states make us think that we are imagining different things, the conscious pain and the brain activity, even though they are one and the same.

This might seem unconvincing; two ways of thinking about the same object are common enough. For example, we can think about the same person as Eric Blair, the Burmese policeman, and as George Orwell, the writer. Yet in these cases any confusion is easily dispelled by evidence that they are the same person. By contrast, phenomenal properties continue to appear distinct from physical properties, even given the persuasive arguments for identifying them, which might seem to argue that they really are distinct.

However, our two ways of thinking about mental states are much more confusing than two names for a person, and this alone can make the states seem distinct even though they are not. Simulating a mental state feels like the state being simulated. Thinking observationally about the corresponding brain state does not. So when we compare these two modes of thought – some grey matter vibrating, versus this feeling – we note that the external mode of thought, by contrast with the simulation, seems to leave out the conscious pain itself. And this makes us conclude that the external mode of thought is therefore about something different, the mere brain state.

If we pause to unpick this last thought, the fallacy becomes apparent. It doesn't follow, from the fact that we don't feel conscious pains when thinking externally about brain states, that these external thoughts are about something different from the conscious pains. After all, we don't in general feel what we are thinking about. So thoughts about brain states are perfectly capable of referring to conscious pains. But we are easily distracted from this. We contrast these external thoughts with the direct simulations once more, and confusedly conclude that thoughts about brain states leave out the conscious feelings, even though in reality the feelings are identical with the brain states.

I call this the "antipathetic fallacy", since the dualist makes the opposite error to Ruskin's poet, by refusing to attribute feelings to brains. It seems to me that it runs deep in our thinking about mind, and accounts for much of the un-

deniable attraction of dualism. Modern physical theory gives us every reason to embrace materialism. But the antipathetic contrast – this feeling, that grey matter – keeps tugging at our thinking and pulling us back towards dualism. We should resist. The temptations of dualism rest on a confusion, and provide no good argument against materialism.

What is more, the identification of the antipathetic fallacy has the further advantage of showing how the materialist should answer Chalmers's more technical arguments. Let me try to convey the gist of these arguments. Chalmers begins, as I said, with the thought that zombies are conceivable. There is nothing contradictory in the idea of a being who is physically just like us, but feels no conscious pain. Modern materialists concede this. They agree that concepts of brain activity are different from the concept of a conscious pain. But, they continue, some such brain concept and the concept of pain might nevertheless refer to the same property, just as the distinct concepts *H₂O* and *water* refer to the same liquid. Chalmers responds that, if this is so, then presumably the brain concept and the pain concept will pick out this supposed common referent by means of different descriptions. The brain concept will identify the common referent in terms of some psychological or physical attribute, while the pain concept will identify it in terms of its phenomenal unpleasantness. But this then gives the game back to dualism once more. For this materialist story itself distinguishes the attribute of phenomenal unpleasantness from any psychological or physical attribute.

This is all cogent enough, if familiar, but it hinges crucially on the assumption that the concept of pain picks out its referent by means of some description. This assumption is contentious, however. The analysis behind the "antipathetic fallacy" shows why. The concept of pain refers by simulation rather than description. We switch on the relevant parts of our brains, and then think "this feeling", rather than picking out pain as something with such-and-such a phenomenal attribute. So Chalmers's final objection to materialism falls away. The difference between the concept of pain and brain concepts is the difference between simulation and external thought, not any difference between associated attributes.

Chalmers briefly alludes to this possibility, but does it scant justice. In a way, he doesn't really have time. Analytic philosophers in the last two decades have analysed the arguments for and against materialism in great detail, and there is now a baroque structure of reasoning surrounding the issue. Chalmers rehearses much of this, often adding his own slant. The arguments I have mentioned just scratch the surface of more than a hundred pages of two-dimensional semantics, epistemic asymmetries and distinctions between different modes of supervenience. By the time Chalmers finally gets to the crucial issue of the structure of the concept of pain, he gives the impression of having lost interest. He only gives it a couple of unsatisfactory pages, plus a cluster of footnotes which betray some residual worries.

Perhaps there is a moral here. Analytic philosophy is sometimes accused of scholasticism. Critics doubt whether its careful distinctions and sophisticated logical techniques ever lead to genuine philosophical progress. I think this charge can be refuted. It would be hard to deny, for example, that recent discussions of materialism have advanced our understanding significantly. At the same time, however, there is a real danger of misplaced technicality. Philosophers can spend so long ironing out the details that they lose track of the main point. It is disappointing, to say the least, to find a talented philosopher like Chalmers mounting so high-powered a dualist

case, yet quite failing to notice that he has left materialism an obvious escape route.

If Chalmers fails to refute materialism, where does this leave the reductive theories of consciousness he criticizes? There is good news and bad news for such theories. The good news is that the "hard problem" probably doesn't need answering after all. The bad news is that consciousness may not be the kind of thing we need a theory of in the first place.

Let me take the good news first. The "hard problem", remember, was to explain the existence of phenomenal properties. Why does the feeling of pain arise when the pain centre in our brain is stimulated? This seems like a good question, but my suspicion is that it only makes sense within dualism. Asking why one thing goes with another presupposes that the two things are distinct. So if materialism is correct, then we should stop posing the hard problem. To go on asking why the feeling of pain arises when your brain is stimulated is like asking, "Why is George Orwell the same person as Eric Blair?" after you learn they are identical.

Plenty of philosophers who count themselves as materialists still think the hard problem poses a genuine challenge. Indeed some of them despair of the ability of humans to solve it. But these philosophers have not taken their materialism sufficiently to heart. They slide back into the antipathetic thought that the cold grey matter leaves the feelings out, and then start looking for an explanation of where the feelings come from. But there is nothing to explain here. That is what the brain state feels like if you are in it.

So it is no failing in reductive theories of consciousness that they fail to solve Chalmers's hard problem. Their only real task is the easier problem of identifying the physical nature of consciousness. The bad news, however, is that even this easier task may be ill-posed. Maybe consciousness is not the kind of phenomenon which can be captured by a reductive scientific theory to start with.

Reductive theories may seem opposed to dualism, but underneath I suspect that many are motivated by it. Even reductionists can be in the grip of antipathetic thinking. They start by thinking of consciousness as some extra non-physical phenomenon, and then try to draw a line around the physical circumstances where it can be found. Once we reject the idea of consciousness as some non-physical inner light, however, it becomes less clear what reductive theories of consciousness are after. The trouble is that the boundaries of consciousness become hazy. Nobody wants to deny, of course, that some states are "consciously like something" for the beings that have them. But unless we think of these states as involving some extra inner light, there is no obvious reason to think that they share any scientifically significant property. Why should there be anything useful to say about what human pains, emotions and visual experiences have in common, not to mention the sensations of chimpanzees, cows and sea cucumbers?

If this is right, then the current enthusiasm for consciousness studies rests on a mistake. It is a discipline without a subject. There is, of course, room for reductive theories of specific psychological abilities, like human vision, or cow learning. But there may be no corresponding theory of consciousness as such. If we take materialism seriously, as we should, then we should stop thinking of consciousness as a distinct phenomenon, a special kind of inner illumination. And once we manage to do this, I suspect that the whole vogue for theories of consciousness will come to seem like a curious fad.

David Papineau is Professor of Philosophy at King's College London and the author of Philosophical Naturalism, 1993.

TWO AWARD-WINNING REFERENCE WORKS PUBLISHED BY
THE BRITISH LIBRARY

**A FOOTBALL COMPENDIUM:
A COMPREHENSIVE GUIDE TO
THE LITERATURE OF
ASSOCIATION FOOTBALL**
compiled by Peter J Seddon
542 pages, cased
ISBN 0 7123 1075 4 £60.00
*Winner of the Library Association
1996 Besterman Medal for an
outstanding bibliography*

**LOCATION REGISTER OF
ENGLISH LITERARY
MANUSCRIPTS AND LETTERS:
18TH AND 19TH CENTURIES**
edited by David C Sutton
1200 pages, 2 vols, cloth
ISBN 0 7123 0388 X £140.00
*Shortlisted for the Library Association
1996 McColvin Medal for an
outstanding reference work*

Available (post-free in UK) from The British Library, Turpin Distribution Services,
Blackhorse Road, Letchworth, Herts SG6 11IN
Telephone 01462-672555 Fax 01462-480947 Email turpin@rsc.org

THE BRITISH LIBRARY
The world's leading resource for scholarship, research and innovation

Source Citation

Papineau, David. "A universe of zombies?" *The Times Literary Supplement*, no. 4864, 21 June 1996, pp. 3+. *The Times Literary Supplement Historical Archive*, link.gale.com/apps/doc/EX1200482985/TLSH?u=tlsacc&sid=bookmark-TLSH&xid=452dcf87. Accessed 25 June 2024.

Gale Document Number:GALE|EX1200482985