

# Consciousness myth

GALEN STRAWSON

---

Tom Stoppard's 'hard problem' may be the hardest there is – but it certainly is not new

---

Published: 25 February 2015

---

mid-1990s, after long neglect.

It's too late to uproot it now. It's spread like Japanese kudzu or Russian ivy. Too many people have a stake in it, including those who believe that they lived through the resurgence (especially the graduate students of the time) and have a place, however modest, among its champions. It soared on a soaring internet whose massively accumulative character then fixed it in place. So it's worth putting it on the record that it's a myth.

In the case of psychology the story of resurgence has some truth. There are doubts about its timing. The distinguished psychologist of memory Endel Tulving places it in the 1980s. "Consciousness has recently again been declared to be the central problem of psychology", he wrote in 1985, citing a number of other authors. The great dam of behaviouristic psychology was cracking and spouting. It was bursting. Even so, there was a further wave of liberation in psychology in the 1990s. Discussion of consciousness regained full respectability after seventy years of marginalization, although there were of course (and still are) a few holdouts.

In the case of philosophy, however, the story of resurgence is simply a myth. There was a small but fashionable group of philosophers of mind who in the 1970s and 80s focused particularly on questions about belief and "intentionality", and had relatively little to say about consciousness. Their intensely parochial outlook may be one of the origins of the myth. But the problem of consciousness, the "hard problem", remained central throughout those years. It never shifted from the heart of the discipline taken as a whole.

It first established itself there (in the modern era) around the time of Descartes's work in the 1640s – for a very specific reason. The rapidly evolving mechanistic corpuscularian theory of matter made it seem clear that matter just wasn't the kind of stuff that could possibly be conscious. Little bits of different shapes and sizes bumping into each other just couldn't do that kind of thing. Leibniz captured the thought in 1714 in a famous image: "*consciousness . . . is inexplicable on mechanical principles*, i.e. by shapes and movements. If we imagine a machine whose structure makes it think, sense, and be conscious, we can conceive of it being enlarged in such a way that we can go inside it like a mill. Suppose we do: visiting its insides, we will never find anything but parts pushing each other – never anything that could explain a conscious state".

This is the so-called "mind–body problem" (the "matter–consciousness problem" is a better name, although there's more to physical reality than matter). It entrenched itself in the seventeenth century, although even then a good

"Many historians of philosophy, with all their intended praise, . . . attribute mere nonsense . . . to past philosophers", as Kant pointed out in 1790. The history of ideas is a zoo – of myths about what happened and what people said. I used to think the mythologizing was a relatively slow process, because the passage of time was needed to blur the past. Twenty years ago, however, an instant myth was born: a myth about a dramatic resurgence of interest in the topic of consciousness in philosophy, in the

number of people, including the deep Hobbes, found no insuperable difficulty in the idea that consciousness was wholly material. It has remained at the centre of philosophical debate ever since. It rocked the second half of the seventeenth century. It consumed the eighteenth, when many, following Descartes's lead, speculated about the possibility of what philosophers now call "zombies", creatures that are not conscious but are outwardly – and perhaps also inwardly – indistinguishable from human beings. In 1755 Charles Bonnet observed that God "could create an automaton that would imitate perfectly all the external and internal actions of man". In 1769, following Locke, he made a nice point against those who resisted materialism on religious grounds: "if someone ever proved that the mind is material, then far from being alarmed, we should have to admire the power that was able to give matter the capacity to think".

The problem of how matter can be conscious remained intractable (a word often used). It roared on into the nineteenth century, even as materialism grew in strength and Darwin, in 1838, followed Georges Cabanis, in 1802, in thinking that consciousness was just "a secretion of brain" (it was a popular image). The problem was regularly announced to be impossibly hard. Increasingly, however, this was taken to be a reason not for doubting that consciousness was material, but rather for doubting our ability to understand the ultimate nature of matter.

This modesty about our powers of understanding was again old – it was as old as it was sensible – but it received powerful new expression. There was a fine cluster in the 1870s, as people speculated about the neural correlates of consciousness. The great German physiologist Emil Du Bois-Reymond delivered his famous verdict on the matter–consciousness problem in 1872: "*Ignoramus et ignorabimus*". We don't know how it is that matter is conscious (although it surely is), and we will never know. In 1874 T. H. Huxley ("On the Hypothesis that Animals Are Automata, and Its History") and W. K. Clifford ("Body and Mind") took it to be far beyond reasonable doubt that consciousness is wholly material – "what I perceive as your brain is really in itself your consciousness, is You", Clifford wrote – in spite of an "enormous gulf" in our understanding. In the same year John Tyndall gave his great "Belfast Address", which contained "that lucky paragraph which has been quoted so often", according to William James in 1890, "that every one knows it by heart":

the passage from the physics of the brain to the corresponding facts of consciousness is unthinkable. Granted that a definite thought and a definite molecular action in the brain occur simultaneously; we do not possess the intellectual organ, nor apparently any rudiment of the organ, which would enable us to pass, by a process of reasoning, from one to the other.

The Belfast Address is long forgotten; the matter–consciousness problem is not. It roared on again, into the twentieth century, and continues unabated in the twenty-first. I'll stick to the last hundred years.

In 1913 the American psychologist J. B. Watson issued a manifesto for behaviourism in psychology. It was at first a purely methodological approach. It didn't say or suggest that consciousness didn't exist. It simply said that you couldn't do proper science with it: the data of consciousness, as delivered by introspective report, weren't susceptible of scientific treatment. So although – of course – consciousness existed, it was better to pass it by in the lab and study wholly publicly observable things like behaviour.

This was a very fruitful move at the time. It's still useful. The trouble started, I'm sorry to say, when some philosophers got their hands on it. They took an admirable method and gamed it into a crazy metaphysics. With a fine fiddle of theoretical bells and whistles, they affirmed that consciousness is really just a matter of behaviour and dispositions to behaviour: nothing more. This was *philosophical* behaviourism (although psychologists were by no means immune). It took off in the 1920s, opening a second front in the debate about consciousness in philosophy and science, even as the original front (the straight-up matter–consciousness problem) remained highly active. In *The Analysis of Mind* (1921), which Watson read in typescript, Bertrand Russell was already worried that behaviourists might go too far. In *Evolutionary Naturalism* (1922), the leading American philosopher R. W. Sellars was alarmed by the behaviourist "tendency either to ignore or to deny consciousness".

Philosophical behaviourists stoutly deny that they deny the existence of consciousness. They say they're simply

giving a new theory of what it is, in saying that it's nothing more than a matter of behaviour and dispositions to behaviour. In fact, of course, to say that consciousness is nothing more than behaviour and dispositions to behaviour is to deny its existence. It's to *eliminate* consciousness – to endorse what duly came to be known as “eliminativism” – and, unsurprisingly, all remotely sensible philosophers and scientists demurred right from the start, among them a top UK team: Russell, G. E. Moore, A. N. Whitehead, Arthur Eddington, C. D. Broad, and Samuel Alexander. In his large and very clever book *The Mind and its Place in Nature* (1925) Broad apologized for bothering to take time to refute philosophical behaviourism before getting down to the real problem of consciousness: “I may be accused of breaking a butterfly on a wheel in this discussion of Behaviourism”. It is, he says, a silly theory, where “by a ‘silly’ theory I mean one which may be held at the time when one is talking or writing professionally, but which only an inmate of a lunatic asylum would think of carrying into daily life”. But “it is important to remember that a theory which is in fact absurd may be accepted by the simple-minded because it is put forward in highly technical terms by learned persons who are themselves too confused to know exactly what they mean”.

Did Broad know what was coming (the “logical” behaviourists were about to weigh in)? Perhaps he did: in his Preface he proposed to “retire to my well-earned bathchair, from which I shall watch with a fatherly eye the philosophic gambols of my younger friends as they dance to the highly syncopated pipings of Herr Wittgenstein’s flute”. The present point, however, is simply that the question of consciousness was central – quite independently of the rising popularity, in the new (i.e. post-1925) quantum mechanics, of claims about the indispensability of reference to consciousness in accounting for quantum-mechanical phenomena. The supposed task was still what it was in Descartes’s time: to “render *intellectually conceivable* the presence of consciousness in the organism”, in R. W. Sellars’s words.

Independently of the behaviourist debate, Russell (*The Analysis of Matter*, 1927), Eddington (*The Nature of the Physical World*, 1928), Whitehead (*Process and Reality*, 1929) and others were pushing forward on the correct philosophical approach. You don’t absurdly deny the existence of consciousness, as some philosophers (e.g. Daniel Dennett) still do today. You don’t – to borrow Broad’s terminology – put forward the silliest view that has ever been held in the whole history of the human race. Instead you raise doubts about how well we know the nature of matter. You think through the point that “Matter”, as Auden remarked in 1940, is, like love, “much / Odder than we thought”. You see that there is – must be – more to matter than physics can tell us. You see that this is in fact a massive understatement. For what does physics tell us about the ultimate intrinsic nature of matter, considered apart from its mathematically expressible structure? Nothing. As Eddington said in 1928, “Trinculo might have been referring to modern physics in the words, ‘This is the tune of our catch, played by the picture of Nobody’”. “If you want a concrete definition of matter”, he said, “it is no use looking to physics.” Physics can’t get at “its inner unget-atable nature”. Russell in 1927 was equally emphatic: “Physics is mathematical, not because we know so much about the physical world, but because we know so little: it is only its mathematical properties that we can discover. For the rest, our knowledge is negative”. This simple point was almost completely forgotten in the 1960s.

On to the 1930s. In his Presidential Address to the American Philosophical Association in 1935, J. B. Pratt observed that “there has never been a time when so much was written in the attempt to solve [the mind–body problem, i.e. the problem of consciousness] or when those who consider this discussion an absurd waste of energy spent so much of their time in trying to prove that no one should spend any time upon it”. Philosophers, psychologists, scientists and theologians were equally engaged on both fronts, the old matter–consciousness front and the new behaviourist front. The debate streamed on into the 1940s. (It was probably in 1944 that Wittgenstein reached the conclusion that a sensation of pain “is not a something, but it’s not a nothing either!”).

All was not lost. In 1948, E. G. Boring, one of the leading “operationist” psychologists in the mid-twentieth century, stood up for the correct common-sense view that “consciousness is what you experience immediately”. In 1950, however, the psychologist and philosophical behaviourist Brian Farrell judged Boring’s claim to be a “comical and pathogenic remark”. The ordinary notion of consciousness, Farrell said, “can be shown to resemble an occult notion like “witchcraft” in a primitive community that is in the process of being acculturated to the West”. Fortunately, he said, science “is getting to the brink of rejecting it, in effect, as ‘unreal’ or ‘non-existent’”.

Not so. In the 1950s the battles over behaviourism were still being fought across science, philosophy, and psychology, fiercely relubricated by two books: Gilbert Ryle's *The Concept of Mind* (1949), which appears to mock ordinary belief in consciousness as belief in a "ghost in the machine", and Wittgenstein's *Philosophical Investigations* (1953). The problem of consciousness remained central. The "ghost in the machine" became famous. Looking back thirty-five years later, Jerry Fodor recalled that "when I was a boy in graduate school [1956–9], the philosophy of mind had two main divisions: the mind/body problem and the problem of other minds". The problem of other minds was discussed mainly in Wittgensteinian terms, but the problem of consciousness lay at its heart, too, and the matter–consciousness problem continued to be debated in its original form.

It got a huge kick of new energy in the second half of the 1950s from U. T. Place's paper "Is Consciousness a Brain-Process?" (1956) and J. J. C. Smart's paper "Sensations and Brain Processes" (1959) – along with Herbert Feigl's long essay "The Mental and The Physical", which was published in 1958 (the same year as Schrödinger's *Mind and Matter*). All three propounded versions of the "identity theory": mind and brain are identical. Mental goings-on are neural goings-on. It was very old news, but it triggered, under the name "Australian materialism", a huge outpouring of discussions of consciousness through the 1960s and beyond. Feigl wrote in 1967 that "the mind–body problems [as opposed to behaviourist issues] are once again in the forefront of highly active and intelligent philosophical discussions". A lot of the participants were in effect committed to some form of eliminativism, i.e. to denying the existence of consciousness, usually in a somewhat covert fashion (Paul Feyerabend and Richard Rorty were more forthright). But this hardly diminished the centrality of the problem. Nor was it undermined by the rapid rise of "functionalism" in the 1960s. It was, on the contrary, increased. For functionalism, as a doctrine in the philosophy of mind, is the view that one can capture the whole nature or essence of certain (or all) types of mental state or occurrence simply by giving an account of their typical or characteristic causes and effects, and it too ultimately amounts to a form of eliminativism about consciousness.

So although the 1970s and 80s – during which functionalism was most fashionable – are the leading candidates for being the lost decades in the discussion of consciousness, they were in fact no such thing. For one of the key issues, and arguably the central issue (this is certainly how it seemed to me when I took up philosophy in 1972 – and certainly no other issue was more central), was the inability of functionalism to give any more satisfactory an account of consciousness than behaviourism. Donald Davidson's "Mental Events" (1970) and Saul Kripke's "Identity and Necessity" (1971) triggered large discussions. Thomas Nagel's famous paper on consciousness, "What is it like to be a bat?", appeared in 1974 and was immediately a focus of discussion (it is one of the most cited papers in all philosophy). Robert Kirk's paper "Zombies v. Materialists" came out in the same year, and talk of zombies was soon ubiquitous. Sydney Shoemaker's ingenious but unsuccessful attempt to find a place for consciousness within functionalism in his paper "Functionalism and Qualia" (1975) ("qualia" is a plural noun used to denote the qualitative characters of conscious experiences such as colour experiences, taste experiences, and so on) further underscored the centrality of the issue, as did Grover Maxwell's "Rigid Designators and Mind–Brain Identity" (1978), in which he re-raised the fading Russellian standard.

So also did all the many bids to "reduce" (in fact eliminate) consciousness, among them Dennett's 1978 "attempt to fit 'consciousness' into information-flow psychology", which, as Ned Block pointed out in the same year, has "the relation to qualia that the U.S. Air Force had to so many Vietnamese villages: he destroys qualia in order to save them". The trouble with functionalism is that it has to allow that unconscious creatures – creatures whom there is "nothing it is like to be" – count as conscious just so long as they function (behave) indistinguishably from conscious creatures. Some muddled souls have thought that this is the issue raised by Alan Turing in his description of the "Imitation Game" in "Computing Machinery and Intelligence" (1950). But when Turing suggests a test for when it would be permissible to describe machines as *thinking*, he explicitly puts aside the question of *consciousness*.

In 1979 Thomas Nagel powerfully restated the difficulty, concluding in his paper "Panpsychism" (which acknowledged the influence of Rebecca Goldstein) that "panpsychism should be added to the current list of mutually incompatible and hopelessly unacceptable solutions to the mind–body problem". John Searle's paper "Minds, Brains, and Programs", containing the famous "Chinese Room Argument", was published in 1980,

illustrating the hollowness of the functionalist idea that a thing incapable of any sort of consciousness could ever really be said to understand anything. It was followed in 1982 by Jackson's "Epiphenomenal Qualia", which told the story of "Mary in the Black and White Room", now known as the "Knowledge Argument" (amusingly dramatized by David Glover in "Brainspotting", a three-part Channel 4 series on consciousness presented by Ken Campbell and broadcast in 1996). Both these papers were focused on the problem of consciousness. Both were among the most discussed papers in philosophy in the 1980s and into the 90s, right up to the mythical point of resurgence and beyond. "Contrary to surface appearances", Searle remarked in 1992, "there really has been only one major topic of discussion in the philosophy of mind for the past fifty years or so, and that is the mind-body problem".

The 1980s also saw the publication of Joseph Levine's much-discussed paper "Materialism and qualia: The explanatory gap" (1983), the same year as Rebecca Goldstein's novel *The Mind-Body Problem*, which contained excellent Russellian insights, and Colin McGinn's *The Subjective View*. Paul Churchland's *Matter and Consciousness*, was published in 1984, Nagel's *The View from Nowhere* in 1986, Michael Lockwood's *Mind, Brain & the Quantum* in 1989. David Rosenthal's "Two Concepts of Consciousness" (1986) also provoked a vast volume of discussion of consciousness – of a refreshingly different kind. It's arguable that it opened a third front in the discussion within analytic philosophy, although it was also part of the attempt to incorporate consciousness into the wholly "naturalistic" account of mind that produced Ruth Millikan's "Biosemantics" (1989).

There was no slackening in the stream of publication. In 1989 McGinn called consciousness "the hard nut of the mind-body problem", in a vivid and influential paper, "Can We Solve the Mind-Body Problem?", in which he gave reasons for thinking that Du Bois-Reymond was right: *ignoramus* and *ignorabimus*. In the same year I called it "the hard part of the mind-body problem", repeating the phrase in *Mental Reality* (1994): consciousness "is the only hard part of the mind-body problem . . . the rest is easy".

We'd reached the mythical watershed of 1995–6, in which philosophers were supposed to wake up to the problem of consciousness, largely on account of Francis Crick's *The Astonishing Hypothesis* and David Chalmers's "Facing up to the Problem of Consciousness", which were both published in 1995, along with Michael Tye's *Ten Problems of Consciousness*, and Chalmers's *The Conscious Mind* in 1996. By this time, however, the 1990s had already produced McGinn's *The Problem of Consciousness* (1991), Dennett's *Consciousness Explained* (1991, sometimes affectionately known as *Consciousness Ignored*), Owen Flanagan's *Consciousness Reconsidered* (1992), and Searle's "The Problem of Consciousness" (1993). Searle's paper was preceded by his book *The Rediscovery of the Mind* (1992), which struck a trenchant blow for common-sense realism about consciousness, especially in Chapter 3, "Breaking the Hold: Silicon brains, conscious robots, and other minds". He was motivated partly by intellectual distress, as he explained in his Introduction: "When I read the standard authors and tried to explain their views to my students, I was appalled to discover that with few exceptions these authors routinely denied what I thought were simple and obvious truths about the mind. It was then, and still is, quite common to deny, implicitly or explicitly, such claims as the following: We all have inner subjective qualitative states of consciousness".

Throughout the 1990s, then, consciousness was a central topic in philosophy of mind. There was a new buzz, but the ratio of occurrences of the words "conscious" and "mind" in the abstracts of all journals covered by *The Philosophers' Index* was effectively the same in the 1960s and 1990s (0.70 and 0.73 respectively), and again in the 1970s and 2000s (0.82 and 0.80 respectively). "Intentional(ity)" had a surge relative to "conscious" in the 1980s and 1990s, but the problem of consciousness remained central. "Conscious" reached its lowest level relative to "mental" in the 1990s (1.1), picked up in the 2000s (1.36), but didn't return to its 1960s and 1970s levels (1.58 and 1.55 respectively).

These are crude measures, but they're significant. There was, certainly, an explosion of interest in the problem of consciousness outside the philosophy world, as Chalmers's use of the phrase "the hard problem" became famous. This was a great thing, but there was no important new idea, as David Papineau observed in his review of *The Conscious Mind* in the *TLS* (June 21, 1996). It's unfortunate that the beautiful history of this debate has been lost (many now write as if the "hard problem" of consciousness was discovered by Chalmers), not just because it's fascinating, but because almost all the best – most vivid, insightful – work lies in the further past. The "astonishing

hypothesis” was the daily bread of the eighteenth-century French materialists and nineteenth-century German materialists. “That matter thinks [is conscious] is a fact”, the Italian Giacomo Leopardi wrote in 1827.

Didn't the 1990s at least witness an increase in the number of philosophers of mind who were prepared to stand up for unequivocal common-sense realism about consciousness? It seems not. The consensus is that such realism continued then, and continues today, to be a minority view among analytic philosophers of mind. How is this possible? Well, as Daniel Kahneman observes, “we know that people can maintain an unshakable faith in any proposition, however absurd, when they are sustained by a community of like-minded believers”. Cicero adds, correctly, that “there is no statement so absurd that no philosopher will make it”.

At the root of the muddle lies an inability to overcome the Very Large Mistake so clearly identified by Eddington and others in the 1920s – not to mention the lovely Irishman John Toland in 1704, Anthony Collins in 1707, Hume in 1739, Joseph Priestley in 1777, and many others. The mistake is *to think we know enough about the nature of physical reality to have any good reason to think that consciousness can't be physical*. It seems to be stamped so deeply in us, by our everyday experience of matter as *lumpen* stuff, that not even appreciation of the extraordinary facts of current physics can weaken its hold. To see through it is a truly revolutionary experience.

It's what Hilary needs to do, in Tom Stoppard's new play *The Hard Problem*. She challenges her amorous tutor Spike to explain consciousness and insists that “when you come right down to it, the body is made of things” – she means physical *things* – “and things don't have thoughts.” There is, however, no good reason to think that this last thing is true, and overwhelming reason to think it's false.

---