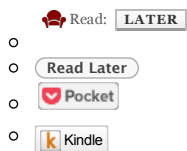


The mental block

Consciousness is the greatest mystery in science. Don't believe the hype: the Hard Problem is here to stay

by **Michael Hanlon**

- Read later or Kindle



Detail from the visualization of the model juvenile rat cortical column, as created by the Blue Brain Project in Lausanne, Switzerland. *Photo courtesy EPFL/Blue Brain Project*

Over there is a bird, in silhouette, standing on a chimney top on the house opposite. It is evening; the sun set about an hour ago and now the sky is an angry, pink-grey, the blating rain of an hour ago threatening to return. The bird, a crow, is proud (I anthropomorphise). He looks cocksure. If it's not a he then I'm a Dutchman. He scans this way and that. From his vantage point he must be able to see Land's End, the nearby ramparts of Cape Cornwall, perhaps the Scillies in the fading light.

What is going on? What is it like to be that bird? Why look this way and that? Why be proud? How can a few ounces of protein, fat, bone and feathers be so sure of itself, as opposed to just being, which is what most matter does?

Old questions, but good ones. Rocks are not proud, stars are not nervous. Look further than my bird and you see a universe of rocks and gas, ice and vacuum. A multiverse, perhaps, of bewildering possibility. From the spatially average vantage point in our little cosmos you would barely, with human eyes alone, be able to see anything at all; perhaps only the grey smudge of a distant galaxy in a void of black ink. Most of what is is hardly there, let alone proud, strutting, cock-of-the-chimney-top on an unseasonably cold Cornish evening.

We live in an odd place and an odd time, amid things that know that they exist and that can reflect upon that, even in the dimmest, most birdlike way. And this needs more explaining than we are at present willing to give it. The question of how the brain produces the feeling of subjective experience, the so-called ‘hard problem’, is a conundrum so intractable that one scientist I know refuses even to discuss it at the dinner table. Another, the British psychologist Stuart Sutherland, declared in 1989 that ‘nothing worth reading has been written on it’. For long periods, it is as if science gives up on the subject in disgust. But the hard problem is back in the news, and a growing number of scientists believe that they have consciousness, if not licked, then at least in their sights.

A triple barrage of neuroscientific, computational and evolutionary artillery promises to reduce the hard problem to a pile of rubble. Today’s consciousness jockeys talk of p-zombies and Global Workspace Theory, mirror neurones, ego tunnels, and attention schemata. They bow before that *deus ex machina* of brain science, the functional magnetic resonance imaging (fMRI) machine. Their work is frequently very impressive and it explains a lot. All the same, it is reasonable to doubt whether it can ever hope to land a blow on the hard problem.

For example, fMRI scanners have shown how people’s brains ‘light up’ when they read certain words or see certain pictures. Scientists in California and elsewhere have used clever algorithms to interpret these brain patterns and recover information about the original stimulus — even to the point of being able to reconstruct pictures that the test subject was looking at. This ‘electronic telepathy’ has been hailed as the ultimate death of privacy (which it might be) and as a window on the conscious mind (which it is not).

The problem is that, even if we know what someone is thinking about, or what they are likely to do, we still don’t know what it’s like to be that person. Hemodynamic changes in your prefrontal cortex might tell me that you are looking at a painting of sunflowers, but then, if I thwacked your shin with a hammer, your screams would tell me you were in pain. Neither lets me know what pain or sunflowers feel like for you, or how those feelings come about. In fact, they don’t even tell us whether you really have feelings at all. One can imagine a creature behaving exactly like a human — walking, talking, running away from danger, mating and telling jokes — with absolutely no internal mental life. Such a creature would be, in the philosophical jargon, a zombie. (Zombies, in their various incarnations, feature a great deal in consciousness arguments.)

Why might an animal need to have experiences (‘qualia’, as they are called by some) rather than merely responses? In this magazine, the American psychologist David Barash summarised some of the current theories. One possibility, he says, is that consciousness evolved to let us to overcome the ‘tyranny of pain’. Primitive organisms might be slaves to their immediate wants, but humans have the capacity to reflect on the significance of their sensations, and therefore to make their decisions with a degree of circumspection. This is all very well, except that there is presumably no pain in the non-conscious world to start with, so it is hard to see how the need to avoid it could have propelled consciousness into existence.

Ray Kurzweil, the Messiah of the Nerds, thinks that in about 20 years or less computers will become conscious and take over the world (Kurzweil now works for Google)

Despite such obstacles, the idea is taking root that consciousness isn’t really mysterious at all; complicated, yes, and far from fully understood, but in the end just another biological process that, with a bit more prodding and poking, will soon go the way of DNA, evolution, the circulation of blood, and the biochemistry of photosynthesis.

Daniel Bor, a cognitive neuroscientist at Sussex University, talks of the ‘neuronal global workspace’, and asserts that consciousness emerges in the ‘prefrontal and parietal cortices’. His work is a refinement of the Global Workspace Theory developed by the Dutch neuroscientist Bernard Baars. In both schemes, the idea is to pair up conscious experiences with neural events, and to give an account of the position that consciousness occupies among the brain’s workings. According to Baars, what we call consciousness is a kind of ‘spotlight of attention’ on the workings of our memory, an inner domain in which we assemble the narrative of our lives. Along somewhat similar lines, we have seen Michael Graziano, of Princeton University, suggesting in this magazine that consciousness evolved as a way for the brain to keep track of its own state of attention, allowing it to make sense of itself and of other brains.

Meanwhile, the IT crowd is getting in on the act. The American futurologist Ray Kurzweil, the Messiah of the Nerds, thinks that in about 20 years or less computers will become conscious and take over the world

(Kurzweil now works for Google). In Lausanne in Switzerland, the neuroscientist Henry Markram has been given several hundred million euros to reverse-engineer first rat then human brains down to the molecular level and duplicate the activities of the neurones in a computer — the so-called Blue Brain project. When I visited Markram's labs a couple of years ago, he was confident that modelling something as sophisticated as a human mind was only a matter of better computers and more money.

Yes, but. Even if Markram's Blue Brain manages to produce fleeting moments of ratty consciousness (which I accept it might), we still wouldn't know how consciousness works. Saying we understand consciousness because this is what it does is like saying we understand how the Starship Enterprise flies between the stars because we know it has a warp drive. We are writing labels, not answers.

So, what can we say? Well, first off, as the philosopher John Searle put it in a TED talk in May this year, the conscious experience is non-negotiable: 'if it consciously seems to you that you are conscious, you are conscious'. That seems hard to argue against. Such experience can, moreover, be extreme. Asked to name the most violent events in nature, you might point to cosmological cataclysms such as the supernova or gamma-ray burster. And yet, these spectacles are just heaps of stuff doing stuff-like things. They do not matter, any more than a boulder rolling down a hill matters — until it hits someone.

Compare a supernova to, say, the mind of a woman about to give birth, or a father who has just lost his child, or a captured spy undergoing torture. These are subjective experiences that are off the scale in terms of importance. 'Yes, yes,' you might say, 'but that sort of thing only matters from the human point of view.' To which I reply: in a universe without witness, what other point of view can there be? The world was simply immaterial until someone came along to perceive it. And morality is both literally and figuratively senseless without consciousness: until we have a perceiving mind, there is no suffering to relieve, no happiness to maximise.

While we are looking at things from this elevated philosophical perspective, it is worth noting that there seems to be rather a limited range of basic options for the nature of consciousness. You might, for example, believe that it is some sort of magical field, a soul, that comes as an addendum to the body, like a satnav machine in a car. This is the traditional 'ghost in the machine' of Cartesian dualism. It is, I would guess, how most people have thought of consciousness for centuries, and how many still do. In scientific circles, however, dualism has become immensely unpopular. The problem is that no one has ever seen this field. How is it generated? More importantly, how does it interact with the 'thinking meat' of the brain? We see no energy transfer. We can detect no soul.

If you don't believe in magical fields, you are not a traditional dualist, and the chances are that you are a materialist of some description. (To be fair, you might hover on the border. David Chalmers, who coined the term 'hard problem' in 1995, thinks that consciousness might be an unexplained property of all organised, information-juggling matter — something he calls 'panprotopsychism'.)

Committed materialists believe that consciousness arises as the result of purely physical processes — neurones and synapses and so forth. But there are further divisions within this camp. Some people accept materialism but think there is something about biological nerve cells that gives them the edge over, say, silicon chips. Others suspect that the sheer weirdness of the quantum realm must have something to do with the hard problem. Apparent 'observer effects', Albert Einstein's 'spooky' action at a distance, hints that a fundamental yet hidden reality underpins our world... Who knows? Maybe that last one is where consciousness lives. Roger Penrose, a physicist at Oxford University, famously thinks that consciousness arises as the result of mysterious quantum effects in brain tissue. He believes, in other words, not in magic fields but in magic meat. So far, the weight of evidence appears to be against him.

Reading these giants of consciousness criticise each other is an instructive experience in itself

The philosopher John Searle does not believe in magic meat but he does think meat is important. He is a biological naturalist who thinks that consciousness emerges from complex neuronal processes that cannot (at present) be modelled in a machine. Then there are those like the Tufts philosopher Daniel Dennett, who says that the mind-body problem is essentially a semantic mistake. Finally, there are the arch-eliminativists who appear to deny the existence of a mental world altogether. Their views are useful but insane.

Time to take stock. Lots of clever people believe these things. Like the religions, they cannot all be right (though they might all be wrong). Reading these giants of consciousness criticise each other is an instructive

experience in itself. When Chalmers aired his ideas in his book *The Conscious Mind* (1996), this philosopher, a professor at both New York University and the Australian National University, was described as ‘absurd’ by John Searle in *The New York Review of Books*. Physicists and chemists do not tend to talk like this.

Even so, let’s say we can make a machine that thinks and feels and enjoys things; imagine it eating a pear or something. If we do not believe in magic fields and magic meat, we must take a functionalist approach. This, on certain plausible assumptions, means our thinking machine can be made of pretty much anything — silicon chips, sure; but also cogwheels and cams, teams of semaphorists, whatever you like. In recent years, engineers have succeeded in building working computers out of Lego, scrap metal, even a model railway set. If the brain is a classical computer — a universal Turing machine, to use the jargon — we could create consciousness just by running the right programme on the 19th-century Analytical Engine of Charles Babbage. And even if the brain isn’t a classical computer, we still have options. However complicated it might be, a brain is presumably just a physical object, and according to the Church-Turing-Deutsch principle of 1985, a quantum computer should be able to simulate any physical process whatsoever, to any level of detail. So all we need to simulate a brain is a quantum computer.

And then what? Then the fun starts. For if a trillion cogs and cams can produce (say) the sensation of eating a pear or of being tickled, then do the cogs all need to be whirling at some particular speed? Do they have to be in the same place at the same time? Could you substitute a given cog for a ‘message’ generated by its virtual-neighbour-cog telling it how many clicks to turn? Is it the cogs, in toto, that are conscious or just their actions? How can any ‘action’ be conscious? The German philosopher Gottfried Leibniz asked most of these questions 300 years ago, and we still haven’t answered a single one of them.

The consensus seems to be that we must run away from too much magic. Daniel Dennett dismisses the idea of ‘qualia’ (perhaps an unfortunately magical-sounding word) altogether. To him, consciousness is simply our word for what it feels like to be a brain. He told me:

We don’t need something weird or an unexplained property of biological [matter] for consciousness any more than we need to posit ‘fictoplasm’ to be the mysterious substance in which Sherlock Holmes and Ebenezer Scrooge find their fictive reality. They are fictions, and hence do not exist ... a neural representation is not a simulacrum of something, made of ‘mental clay’; it is a representation made of ... well, patterns of spike trains in neuronal axons and the like.

David Chalmers says that it is quite possible for a mind to be disconnected from space and time, but he insists that you do at least need the cogwheels. He says: ‘I’m sympathetic with the idea that consciousness arises from cogwheel structure. In principle it could be delocalised and really slow. But I think you need genuine causal connections among the parts, with genuine dynamic structure.’

As to where the qualia ‘happen’, the answer could be ‘nowhere and nowhen’. If we do not believe in magic forcefields, but do believe that a conscious event, a quale, can do stuff, then we have a problem (in addition to the problem of explaining the quale in the first place). As David Chalmers says, ‘the problem of how qualia causally affect the physical world remains pressing... with no easy answer in sight’. It is very hard to see how a mind generated by whirring cogs can affect the whirring of those cogs in turn.

Nearly a quarter of a century ago, Daniel Dennett wrote that: ‘Human consciousness is just about the last surviving mystery.’ A few years later, Chalmers added: ‘[It] may be the largest outstanding obstacle in our quest for a scientific understanding of the universe.’ They were right then and, despite the tremendous scientific advances since, they are still right today. I do not think that the evolutionary ‘explanations’ for consciousness that are currently doing the rounds are going to get us anywhere. These explanations do not address the hard problem itself, but merely the ‘easy’ problems that orbit it like a swarm of planets around a star. The hard problem’s fascination is that it has, to date, completely and utterly defeated science. Nothing else is like it. We know how genes work, we have (probably) found the Higgs Boson; but we understand the weather on Jupiter better than we understand what is going on in our own heads. This is remarkable.

Consciousness is in fact so weird, and so poorly understood, that we may permit ourselves the sort of wild speculation that would be risible in other fields. We can ask, for instance, if our increasingly puzzling failure to detect intelligent alien life might have any bearing on the matter. We can speculate that it is consciousness that gives rise to the physical world rather than the other way round. The 20th-century British physicist James Hopwood Jeans speculated that the universe might be ‘more like a great thought than like a great machine.’ Idealist notions keep creeping into modern physics, linking the idea that the mind of the observer is somehow fundamental in quantum measurements and the strange, seemingly subjective nature of time itself, as

pondered by the British physicist Julian Barbour. Once you have accepted that feelings and experiences can be quite independent of time and space (those causally connected but delocalised cogwheels), you might take a look at your assumptions about what, where and when you are with a little reeling disquiet.

I don't know. No one does. And I think it is possible that, compared with the hard problem, the rest of science is a sideshow. Until we get a grip on our own minds, our grip on anything else could be suspect. It's hard, but we shouldn't stop trying. The head of that bird on the rooftop contains more mystery than will be uncovered by our biggest telescopes or atom smashers. The hard problem is still the toughest kid on the block.

Correction, 10 Oct 2013: The original version of this article stated that Charles Babbage's Difference Engine would have been Turing-complete. In fact it was Babbage's Analytical Engine that had this distinction. We regret the error.



aeon
FILM

[Visit our new film channel](#)



Follow My Steps

A 13-year-old with muscular dystrophy finds a big brother to help him through the trials of adolescence

Andrew Hida

16 minutes

Published on 9 October 2013