'The canniest science writer around. He writes, above all, with attitude.'
INDEPENDENT

AT THE EDGE OF UNCERTAINTY

MICHAEL BROOKS

11 Discoveries Taking Science by Surprise

Author of the bestselling *13 Things That Don't Make Sense*

First published in Great Britain in 2014 by
PROFILE BOOKS LTD
3A Exmouth House
Pine Street
London EC1R OJH
www.profilebooks.com

Extract from ‘Little Gidding’ taken from *Four Quartets*, by T. S. Eliot (Faber & Faber, 2001). Reproduced by kind permission of Faber & Faber.

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

eISBN 978 1 84765 966 8

CONTENTS

ACKNOWLEDGEMENTS

A project like this always involves battling with uncertainty, and I am grateful to many people for helping to shape and inform the content of this book. The enthusiasm of Andrew Franklin and his colleagues at Profile provided steadfast support through the process of pulling these ideas and pages together. My agent, Caroline Dawnay at United Agents, has been an ever-present source of helpful advice and direction. I am indebted to the many people who responded to my pestering emails, giving freely of their time to read through various parts of the manuscript and point out where there was room for improvement. Particularly worthy of thanks are Giovanella Baggio, Martin Bobrow, Daniel Bor, Carl Gibson, Neal Halsey, Grazyna Jasienska, Rajeev Krishnadas, Karen Lillycrop, Toby Ord, Krill Rossiianov, Rudolph Schild, Jack Copeland, Hava Siegelmann and Vlatko Vedral. They helped limit my uncertainties, but any remaining errors are mine.

As ever, I must thank those closest to me for coping with more months in the shadow of my inattention. Phillippa, Millie and Zachary, our family is a deep well of certainty; thank you.

INTRODUCTION

Daring ideas are like chessmen moved forward. They may be beaten, but they may start a winning game.

Johann Wolfgang von Goethe

You might think it's hard to take science by surprise. After all, aren't scientists the clever ones, the know-it-alls? Aren't they revered as the people with answers to every question?

It's certainly true that science has made extraordinary inroads into discovering how the universe and everything within it ticks along. Science has been successful for the most part in explaining why things are as they are. But in the process they have also discovered the broad horizon of their ignorance.

That is not a problem; on the contrary, it is an enormous gain. In science, ignorance is not something to be ashamed of, something to hide, but something to acknowledge and explore. Just as the tide's ebb and flow created the perfect conditions for life to arise at the edge of Earth's oceans, the place where certainty gives way to uncertainty – the shoreline of our ignorance – is fertile ground indeed.

In much of science, the parts we know well, there is relatively little to be gained. Here, further up the beach, we might determine a constant to another decimal place; there, we seek to make a slightly more accurate measurement of the time it takes for a signal to travel between neurons in the brain. We find a catalyst that will make a chemical reaction happen a little more quickly or efficiently. We discover another distant star to enter into our catalogues, and so on. Such incremental gains are always there for the taking, pebbles to be turned over and inspected. These advances are added to the canon of science, but they don't change anything – not really. That is why they don't make frontpage news. Newton was too humble when he wrote about his life's work shortly before his death. He said, 'I was like a boy playing on the sea-shore, and diverting myself now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.' It's not true: much of what he did was reaching into the murky water and pulling out surprising new truths.

Many have followed in his footsteps, moving out of the safe zone, venturing beyond the very limits of our knowledge and peering into the gloom until they could make out the vague shape of something intriguing. Then, grabbing all the tools at their disposal, they plunged into the water, intent on bringing that shadowy form back on to dry land.

It is a dangerous thing to do. Here at the edge of uncertainty we have discovered shocking things – things that have made some scientists beat a hasty retreat. It was here, for instance, that Henri Poincaré discovered that a resolution to certain anomalies in electromagnetic theory would require rethinking the nature of time. Poincaré was too perturbed by the discovery to press on; it was left to Albert Einstein to venture into dark waters and hunt out the special theory of relativity. The astronomer Arthur Eddington had once done some work that suggested the existence of black holes, but he hated the implications: that there were rips in the fabric of the universe. So when Subrahmanyan Chandrasekhar confirmed the suggestion with a mathematical proof, Eddington railed against it, and made Chandrasekhar's life a misery. Neuroscientist Benjamin Libet was another fugitive from unwelcome truth; when he performed an experiment showing that humans lacked free will, he dedicated the rest of his life to proving himself wrong. Good science – important science – can be as unnerving as it is enlightening.

Sometimes work at the edge of uncertainty is without tangible fruit: it simply uncovers our ignorance. From time to time, for example, we will discover that our previous scientific understanding was built on a flimsy foundation and must be urgently shored up – or even abandoned. This is not the disaster it might seem, because science is fickle: it reserves the right to change its mind. Some scientists might make definitive statements, but others must then take on the task of trying to undermine them. Very often they succeed: new experiments, new thoughts and new discoveries turn our thinking on its head, reverse a trend, expose the flaws in previous experiments, or poke holes in a celebrated scientist's thinking. The initial result is usually panic or denial, anger or derision – often all of the above. Eventually, though, after months, a year, a decade or a century, there is resigned acceptance of the new. Until, that is, someone dares to take in the view from the new edge of uncertainty. That novel perspective inevitably leads to further revisions and revolution. 'Everything we know is only some kind of approximation,' Richard Feynman once said. 'Therefore, things must be learned only to be unlearned again or, more likely, to be corrected.' This is where Galileo, Newton, Darwin and Einstein did their work. All the revolutionaries have been challenged, accepted, then challenged again. As George Bernard Shaw put it, 'All great truths begin as blasphemies.'

Where science does have a problem is in the fact that our collective memories are so short. Once that resigned acceptance of a discovery comes, we forget that there was once such a kerfuffle. We act as if this truth were always with us, that it is self-evident. We forget the decades of persecution someone endured in order to shepherd us to the view we would now die to defend. And so we become comfortable – so comfortable that we will wantonly persecute the man or woman who comes to disturb our peaceful state. Take the atom, for example. No one now denies its existence, and it seems impossible that anyone thought it a pointless fiction. The atom is part of our world-view, part of our language, part of our collective history. But it wasn't always thus, as the tragic story of the Austrian physicist Ludwig Boltzmann shows so clearly.

These days Boltzmann would almost certainly be diagnosed as having bipolar disorder. His moods swung between elation and deep depression. When up, he was convivial – his students loved him and lectures at the University of Vienna were sometimes attended by so many people that the overspill ran into the corridor and down the stairs from the lecture hall. When down, a phase often triggered by the rejection of his peers, his moods were very dark indeed. In 1900, for instance, after an argument with a member of his department, Boltzmann tried to take his own life.

The rancour was always over the existence of atoms. Boltzmann was convinced that they existed in some form or other; most of his colleagues, some of the most powerful men of science at the time, were convinced they did not. Though the notion of atoms seems rather obvious to us now, and might well have seemed obvious to physicists hundreds of years earlier, many of Boltzmann's contemporaries were obsessed by the nebulous concept of energy. The industrial revolution had raised energy, in their minds, to a position where it became the fundamental component of reality. They believed that the new science of thermodynamics, which had been constructed to further the gains of the industrial revolution from its roaring heat engines, provided reality's rules.

Boltzmann spent the latter part of his working life countering this view. He constructed intricate arguments which proved that the mechanical motion of atoms was the fundamental driving force of gases as they heated, expanded, cooled and contracted. The theory was statistical, not absolute: though individual atoms followed simple rules, together they would create a variety of observable outcomes. Some outcomes were more likely – some much more likely – than others, providing an explanation for the observed phenomena. It was an unpopular notion, with the most popular physicists of the time railing against it. Chief in opposition was Ernst Mach, who admitted that atoms could be a useful crutch for thinking about reality but nothing more: the atom, he said, 'must remain a tool for representing phenomena'.

Boltzmann defended his position with gusto, but was undoubtedly worn down by the fights and his opponents' air of indifference. In the middle of one debate, he recalled, 'Mach spoke out from the group and laconically said: "I don't believe that atoms exist." This sentence went round and round in my head.'

Boltzmann's head had never been a terribly stable place, and the years of bitter feuds over the atom were a further strain. Eventually Boltzmann decided to end the agony for good. In 1906, while his wife and daughter swam nearby in the blue waters of the Bay of Duina near Trieste, Boltzmann hanged himself. His daughter, sent back to check on her father, found his body hanging by a short cord from a window casement. For the rest of her life, she never once spoke of what she had seen.

The singular literary character Lemony Snicket seems to understand the predicament Boltzmann found himself in. 'It is very unnerving to be proven wrong,' he says in *The Reptile Room*, 'particularly when you are really right and the person who is really wrong is proving you wrong and proving himself, wrongly, right.'

Such is the eternal dilemma of science: it's not always clear who is right, and the truth sometimes emerges too late for its champion to enjoy the victory. We don't know that Boltzmann's suicide resulted directly from the opprobrium raised against him, but we do know he played a vital – if tragic – role in securing a victory for a new and better understanding of reality. Within a few years of his death, observations of pollen and dust grains being knocked about at random by invisible entities led to the acceptance of Boltzmann's atom.

Historians of science know this well, but historical examples are of limited use. 'To imagine that turmoil is in the past and somehow we are now in a more stable time seems to be a psychological need,' the geologist Eldridge Moores once said. He was talking about wishful thinking concerning the stability of the ground beneath our feet, but he might have been talking about science. It's somehow easier to marvel at the fossils – to enjoy these stories of science's evolution – than to accept that things might still be evolving, that there is still an edge of uncertainty over which we can peer.

That is what this book has to offer: a glimpse of today's edge. We should be glad that there is one. After all, we don't have forever to mess about dotting the i's and crossing the t's of science. According to biologists John Lawton and Robert May, the fossil record tells us that a mammalian species lasts a million years on average. We have already been around for one-fifth of that time, and only recently started to study the world around us using what we now recognise as scientific techniques. Maybe the discoveries we make at the edge of uncertainty will help our species be the first to last forever – perhaps what we are studying here is the key to eternity. We certainly won't find the secrets of survival in the extra decimal place of something we know too well already.

We must go down the beach and peer into the dark waters. Extraordinary things are waiting to be discovered. They are almost certainly too extraordinary for us to cope with just yet. But we are becoming accustomed to the murky view, and a few hints and teases are beginning to resolve before our eyes. It is those hints and teases that we are about to explore; the chapters of this book describe some of the dangerous frontiers of science today. They are things we see most dimly, and yet feel most drawn to because of their potential to transform the way we see ourselves and the way we live out our existence.

We start with an acknowledgement of what is perhaps our greatest scientific weakness: the human brain. These few pounds of wobbly biology are our only means of understanding the universe, and yet we don't even know what it means to 'understand'. We think, therefore (we tell ourselves) we are: we have our sense of self. From there we extrapolate to see our selves playing a role in a grand cosmic theatre.

In many ways it is a remarkable, self-aggrandising world-view. After all, as we will see, many other species of animal – species we don't see as struggling with existential questions – are remarkably similar to us. Our discoveries about the abilities of non-human animals has not knocked us off the pinnacle of creation, but it has brought many of our fellow creatures up to join us. These days we highlight the similarities, not the differences. One of the consequences is a movement to suggest we can merge human and non-human animals for medical purposes. It's something we are already doing with the creation of chimeras, though this is still at the edge of uncertainty and only slowly emerging from the scientific (and ethical) darkness.

We are also discovering that we have been transgressing a couple of other ethical boundaries. Were the geneticist Jacob

Bronowski, the man who said that knowledge is 'personal and responsible', still with us today, he wouldn't have hesitated to suggest that our emerging understanding of epigenetics necessitates a headlong chase into the unknown. Epigenetics describes how genes work differently in our bodies after assault and insult from factors generally associated with poverty, deprivation and pollution. The effects are problematic and long-lasting, sometimes cascading down the generations. We are only just beginning to realise how personal biology is – and how responsible our response should be.

The same can be said of our discoveries about the role of gender in medicine: somehow we have been unconscionably crude in our medical approach to human beings. Did we really think that, the obvious points aside, gender makes no difference? We don't any more. Perhaps more forgivable, given the paucity of our understanding of the brain, was our neglect of the mind's power in the body. Here too, though, we are slowly undoing the cynicism of those who were happy to dress up their ignorance as knowledge and understanding.

If only we could do the same with those who tout quantum theory as the key to health. There is an undoubted appeal in calls to employ 'the power of quantum healing to transcend disease and aging', as the mystic Deepak Chopra would have us do. However, it is a mirage in the desert. The truth is, we are only just learning how to make the most tentative investigations of the role quantum physics plays in biology. It does seem that there are areas where nature has exploited the strange rules that govern atoms and molecules to create new opportunities for life to blossom and flourish in adversity. But here, at this boundary between life and the stuff of the cosmos, we truly jump into the deep waters.

As we bring together the sum of our experience in mathematics and physics, in experiment and theory, we can make the tentative suggestion that the universe is a computer, with our thoughts and actions acting as the programs whose instructions create what our brains (our poorly understood brains, remember) interpret as reality. Is this as delusional as Newton's insistence on the 'clockwork heavens' – an interpretation based on the technology of his time? Perhaps. After all, the computer is only a few decades old, and its inventor, Alan Turing, did see another kind of computing machine beyond the one that is familiar to us. Perhaps the hypercomputer will be a better guide to reality.

Not that we are finished with the reality we already know. There are those who would like to close the book on our story of how the cosmos came to be, but others are resisting. There are too many holes in the Big Bang story, and too many places where we have plugs that might fit – or at least know how we might fashion them to fit. It may be that, when we have patched up the history of the universe, there is more patch than fabric and we need to start again. We are already starting again with one of its fundamental constituents: the flow of time, it seems, is nothing more than an illusion. The tick-tock of passing moments is all in our minds, the physicists suggest.

In many ways, it would be easier to ignore all this, to go back up the beach to where the unfinished i's and t's are waiting. After all, we are simple creatures, easily fooled by our senses, our inner logic and our desire to bring simplicity to our interactions with the world. These difficult topics expose our weaknesses and leave us open to failure. Making sense of them is hard.

The beauty of human beings, though, is that we are fierce and indefatigable. We have shown ourselves determined to grapple with the universe around us until it surrenders its secrets to our inquiries. That is why we go to the edge of uncertainty: to quest, and question, and fight with ourselves and others until we have an answer. Then, aware that we have brushed against other questions and surprises, we stow our new discoveries safely, and dive back into the dark waters to wrestle more things into the light. We have been doing it for centuries, and we can only hope we will be doing it for centuries to come. This is, after all, the best thing humans have ever done.

This is how those mysterious and powerful brains compel us to behave: they endow us with the curiosity, the bravery and the tenacity to hunt out the truth as best we can. It's not an easy way to live. By the end of this journey to the frontiers of human certainty and beyond, your brain will feel battered and bruised. But it will also cry out for more. Adventuring is addictive. You have been warned.

1

TRIUMPH OF THE ZOMBIE KILLERS

The science of consciousness has risen from the grave

We have been to the moon, we have charted the depths of the ocean and the heart of the atom, but we have a fear of looking inward to ourselves because we sense that is where all the contradictions flow together.

Terence McKenna

To the audience's utter delight, Gustav Kuhn is performing magic tricks. He makes ping-pong balls disappear and reappear in ridiculous places. Then he explains how he did it. 'It's simple misdirection,' he says. 'I manipulate your attention by moving my hand in certain directions; you can't help but follow it with your eyes, which gives me the chance to…' He turns his head, and our gaze follows. The ball is back in his hand. We can't help but applaud.

It's unusual for an audience to be clapping this early into a scientific talk. Usually there's a smattering of applause at the end – often a manifestation of relief that it's finally over. But here at the sixteenth meeting of the Association for the Scientific Study of Consciousness the audience is enthralled from the start.

Kuhn thinks there should be a science of magic. The effects he and other magicians create are robust, significant, replicable and, above all, useful, he says – just like good scientific results, in other words. He and his co-presenter, Ronald Rensink, another magician–scientist, think that studying what magicians do can teach us about perception and cognition (and deception), how children develop an understanding of what is possible and impossible, why magical beliefs persist and what happens when the brain develops in unexpected ways. A study of magic could help us develop new tricks for engaging and interacting with people and technology and find new angles on problem-solving. And most important of all, it might give us a window on what it means to be conscious.

Studying consciousness used to be considered the ultimate waste of time. It is, after all, a subjective phenomenon, and thus unlike anything else in science. How can I study someone else's consciousness when I have to rely on their reports? How can I study my own, when I can't get any distance from it? Somehow, that spongy matter inside my skull creates something we call consciousness, but if I probe it, I disturb it. We don't have the means to keep a brain alive outside the skull and, even if we did, would we expect to dissect a brain and find its consciousness?

In 1994 philosopher David Chalmers coined a phrase about consciousness that has become a millstone or a mantra, depending on your point of view: 'The Hard Problem'. Consciousness 'escapes the net of reductive explanation,' Chalmers says. 'No explanation given wholly in physical terms can ever account for the emergence of conscious experience.' In other words, consciousness can't be explained by reverse-engineering the brain. You can't build a brain and expect to trace where its consciousness comes from. Consciousness is different in character from the set of all physical facts – it stands apart. That's why, he said, it is possible we are surrounded by undetectable zombies.

Any number of movies have described the onset of a zombie apocalypse. In not one of them has the hero used sleight-of-hand tricks to give their loved ones time to get away. That might seem like a banal observation, but it raises an interesting question about the nature of consciousness – and Chalmers's argument. Would a zombie be amazed and distracted by Kuhn's conjuring? What do zombies make of magic?

To be fair to Chalmers, he wasn't talking about the familiar, flesh-hanging-off, undead, food-obsessed zombies of science fiction. After all, they're easy to spot, with their lumbering gait, their insensitivity to pain or injury, their inability to communicate with or relate to others, and their glassy-eyed stare. What we're talking about is the perfect copy of a normal human, one that, from the outside, looks no different to you or me. This zombie walks normally; it can hold a conversation. It will even tell you it is feeling something. But the first question you have to ask yourself is how you could tell if it is telling the truth. You couldn't.

You can say exactly the same about your work colleagues. You, as Descartes pointed out, know you are conscious – '*cogito, ergo sum*', I think, therefore I am. But how do you know anyone else is conscious? All you have to go on is the fact that they appear to be the same as you. They react to stimuli such as a punch in the arm in the same way you would. Ask them a question and they respond in reasonable ways, and in a reasonable time. But if you ask them what they are experiencing, you have no way of knowing if they are just telling you what they think you expect them to say. They might not feel anything – they might just know what a human being is expected to be feeling in that situation, and report that.

This is the zombie hypothesis: that everyone around you might lack any self-awareness, any sense of self, and you wouldn't necessarily know it. Bringing it closer to home, imagine a version of you that is exactly like you, physically and mentally, so they look, act and speak like you, even thinking like you to give the same answers that you would to any question someone cared to pose. The difference is that this version of you has no awareness of themselves; they are, effectively, an automaton.

The fact that you can imagine this, Chalmers says, means it is theoretically possible. And so, he argues, consciousness must be something extra and above the physical material and processes of our brains, something that sits on top of our sensory perceptions, our reactions to them and our reporting of them.

That 'something' makes us more than a zombie. This difference, we could say, defines consciousness. It's that quality that gives us a sense of self, of what we are feeling, of introspection, examining and questioning our place in the world. It is, perhaps, what makes us amazed and entertained by magic tricks. It is what makes us laugh and cry. It is, you might say, what makes us human. Philosophers have longed to distil this essence of self-awareness for centuries. The exciting thing is that science is now, finally, giving us ways to probe the issue that involve more tricks than just thinking about it. And it appears that our scientific insights have killed the zombie. We can stand astride its corpse and declare that we will win in the end because we now see that consciousness must have a physical root and, consequently, will indeed succumb to science.

All of the work scientists have done on consciousness so far has led us to a handful of models that seek to exemplify what is going on inside our heads. Two are considered most promising. One is the global neuronal workspace theory, a combination of insights from psychology and neuroscience. It suggests that all the inputs from the outside world – touch, taste, vision, hearing and so on – are first processed unconsciously. Very few of these inputs will get your attention; this only happens when there is enough subconscious processing going on to trip a switch that activates the areas of the brain concerned with conscious processing. Neuroscientist Daniel Bor describes it as ‘a spotlight on a stage, or scribbles on a general-purpose cognitive white board’. Put simply, it’s putting our short-term working memory to use – although those memories last only a couple of seconds, it’s long enough to draw on them when necessary.

Its chief competitor is known as information integration theory. This model puts consciousness into the language and framing of information theory, creating datasets that add up to more than the sum of their parts. Its originator is Italian psychiatrist and sleep researcher Giulio Tononi. He is a controversial figure in many ways – though his theory is barely on its feet, he has declared that it could lead to a universal consciousness meter that would measure the ‘level’ of consciousness of anything from a worm to a computer network. However, information integration theory is about the whole network of neurons, and makes no attempt to explain what is going on in the individual physical structures of the brain. That means it doesn't much lend itself to the kinds of simple experiments used to test the global workspace theory. That said, it has some heavyweight fans. 'It's the only really promising fundamental theory of consciousness,' Christof Koch told *New York Times* writer Carl Zimmer.

In the end, though, we have to admit that decades of development have resulted in theories of consciousness that are still somehow unsatisfying. Psychologists and neuroscientists are, in many ways, like Darwin aboard the *Beagle*: they are still gathering specimens and making observations of interesting things done by the brain. They haven't yet, if we're being honest, got very far in pulling it all together into a coherent theory, a simple idea that explains the subjective experience of being conscious of what is around us, of thinking about things, of how the stuff of our brains creates a different experience from the zombie's existence without awareness. And that is exactly what led researchers to kill the zombie.

The leader of the zombie-hunters is unquestionably Tufts University philosopher Daniel Dennett. His strategy is remarkably simple. Perhaps there is, he suggests, no such thing as consciousness, that this ongoing awareness and sense of thinking about the world is actually an illusion. Perhaps our brains are fooling us into thinking there's some overarching narrative to our existence.

In 1991 Dennett published a book with an audacious title. *Consciousness Explained* was greeted with charges of hubris, but perhaps the detractors should have waited. In the book's 'Appendix for Scientists', Dennett made a prediction that, if his theory was right, we should be blind to many subtle changes in our environment. Change-blindness would exist, he said, because the conscious visual experience is not a true reflection of what is actually in front of individuals.

Dennett's idea is similar to the premise of the movie *The Matrix*, where humans have a conception of reality that is actually a carefully stitched simulation fed directly into their brains by a race of machines. In Dennett's view, there are no machines, only the brain. But, just as the machines' simulation sometimes has glitches, if we look carefully enough at our world, we'll see the brain's stitches. It turns out he was right.

Ronald Rensink has done a lot of the work to prove Dennett's hypothesis. He has carried out a series of experiments that show people missing seemingly obvious things right in front of their eyes. To understand why, we can start with the issue of foveal saccades.

The evolution of the eye and the visual processing system has had to cope with a number of efficiency measures, but perhaps the most remarkable one is that, even without taking blinking into account, for around four hours of every waking day there is no visual information being processed by your brain. That's because your retina takes in a full image of the world on a patch of densely packed photoreceptor cells that is about one millimetre in diameter. This is the fovea centralis, which records detail and colour from the world around you. The thing is, it only takes that reading from an area that's about the size of your thumbnail held at arm's length. Your vision captures everything else in front of you at that moment at a much lower resolution, and in monochrome. Go ten degrees off centre and you're capturing about 20 per cent of that maximum amount of visual information. In other words, most of what you see is recorded in a blurred black-and-white image.

The reason you're not aware of having such a 'lo-fi' view of the world is because your eye is constantly flitting about, capturing as much of the visual field as possible on the fovea's receptors. Roughly three times a second, for about 200 milliseconds each time, you record a high resolution image, and then your eye moves again. In between these saccades, or jerky movements, your brain turns off in order to prevent you registering the blurred image of the movement. In a paper published in *Trends in Cognitive Sciences*, David Melcher and Carol Colby showed that spending 100 milliseconds 'offline' roughly 150,000 times per day adds up to four hours of blindness. You don't notice it because your brain stitches together the processing it *has* done, creating the illusion of seamless visual perception. But that's nothing compared to the illusions exposed by change-blindness researchers.

Working with his colleague Daniel Simons, Dennett has performed some stunning (and hugely entertaining) experiments demonstrating how the smooth flow of our visual neurochemical, and neurophysiological substrates of conscious states along with the capacity to exhibit intentional behaviours… humans are not unique in possessing the neurological substrates that generate consciousness. Non-human animals, including all mammals and birds, and many other creatures, including octopuses, also possess these neurological substrates.'

They signed their written deposition, slightly oddly, in the presence of Stephen Hawking. But perhaps it was a good choice. Who would claim that Hawking, a brilliant cosmologist, isn't conscious? He is undoubtedly very much aware of his surroundings, capable of feelings such as joy and sorrow, and a cogent and fearsome thinker. Strip him of the technology that allows him to communicate, however, and of the human carers that meet his physical needs, and it might be possible to plausibly deny his consciousness.

That is why this area of research is so important. An understanding of consciousness is key to relating properly to animals, but it will also help us face our problems with the biggest human dilemma: death.

In 1985, when her husband was hospitalised with pneumonia, doctors asked Jane Hawking if she would like his life-support system turned off. Stephen Hawking was not yet a celebrity figure – *A Brief History of Time* was still an unpublished manuscript – and his diagnosis of motor neurone disease had robbed him of any significant life expectancy. He was in a drug-induced coma; the doctors were open to Jane ending it there.

Just occasionally, our medical skill, our ability to keep people alive, can be seen as a modern curse. Death already stalks us through life. We are among what is a very small number of animals who live with an awareness that we will one day die. It's made even more difficult when we are put in control of the timing. Perhaps that is why we find the coma patient – the seemingly dead living among us – so hard to handle. In the presence of a patient lost to a coma, unreachable and unable to reach out with any form of communication or gesticulation, we are reduced to our own peculiar paralysis. We don't know what is the 'right' thing to do.

Jane Hawking said no. Others haven't, though. Let's hope they haven't heard about Adrian Owen's recent discovery that supposedly disastrous brain damage, even when it results in an apparently 'permanent' vegetative state, can occasionally leave a patient in a conscious state.

Owen's team of researchers assessed a man who, while meeting all internationally agreed criteria for being in a vegetative state, was able to generate 200 responses to direct commands. He couldn't communicate with the outside world, but Owen's electroencephalogram (EEG) reading of the electrical activity in his brain made it clear that he was responding just as you or I would if we happened to be bound and gagged. 'He's probably as conscious as you or I,' is how Owen put it to *New Scientist* reporter Chelsea Whyte.

It's worth noting that this was just one patient in a study of nineteen; only three of those vegetative patients showed reliable signs of consciousness – and researchers continue to debate what constitutes a reliable response – and, of course, what constitutes a conscious response. Nonetheless, there is clearly something worth exploring here. Then there is the extraordinary discovery that coma patients show an emotional response to music that is played to them. Again, they can't show it outwardly, but their heart rate changes in exactly the way yours or mine would. We must be cautious about false inference – we don't know that the heartbeat change necessarily means these patients have any degree of consciousness – but it's another area ripe for exploration. If these people are really just 'locked in', that changes the way we must treat them. Owen suggests the first thing to find out is whether they are in pain. Then we must use technology to allow them to communicate. Brain–machine interfaces are becoming ever more adept at reading and interpreting brain activity; we might be only years away from holding a meaningful conversation with a coma patient. After all, even though he wasn't in a coma, no one should have been able to talk meaningfully to Carissa Philippi's patient, a man known as R.

R was a fifty-seven-year-old college-educated man whose brain had been severely damaged by an episode of herpes simplex encephalitis. The virus had destroyed his insular cortex, anterior cingulate cortex and medial prefrontal cortex. Those three areas of the brain are vital to self-awareness: according to the experts, R should have been a zombie.

He clearly wasn't. When the researchers asked him for his definition of consciousness, R replied with a cogent take on the issue: that it was the body's 'awareness and reaction to what's going on in the environment around it'. The researchers asked him, 'Do you think the sense of self is like a concept?' 'Yes,' he said, 'it's an idea in your brain.' Despite lacking the brain material thought essential for consciousness, R certainly didn't seem like a zombie.

R can't taste or smell, and he suffers from severe memory loss. However, he's not in bad shape, all things considered. The Iowa researchers carried out a series of tests and imaging studies on him. They found that his intelligence is in the normal range, and he recognises himself in the mirror. One of the researchers surreptitiously put black eye-shadow on R's nose, and he showed no hesitation in wiping it off when presented with his reflection fifteen minutes later (that was sufficient time for him to forget completely the researcher's contact with his face). He has a sense of agency over his actions: he knows what effects he causes. He can't, for instance, tickle himself; no one with a sense of their own agency can. When asked to do a little introspection, he reports he has a sense of himself. The researchers' conclusion is stark and straightforward: we have to stop thinking about consciousness as something that sits in a single 'higher', more evolved part of the brain, they say. 'R is a conscious, self-aware, and sentient human being,' the researchers declare; self-awareness, they conclude, 'is likely to emerge from distributed interactions among networks of brain regions.'

In some ways, that's not surprising. We know that the prefrontal parietal network, which seems to play a central role in conscious processing, contains the brain structures that have the most connections to other regions of the brain. Consciousness is, in that way, rather like the Internet. It, too, involves widely distributed processing, with no one place in particular tied to any one task. That is not to say there aren't particular tasks carried out at particular places, but, as R shows, the brain can adapt to injury – and so can the Internet. People have tried to turn it off by destroying parts of it – there was a huge attack in 2007, for instance, which aimed to overwhelm the thirteen servers that carry vital data. Two servers went down, but the other eleven carried on with business as usual. Having survived the attack, the system has been modified so that the servers also act as mirrors for each other: if one goes down, another takes over its role.

There are those who believe that the Internet could eventually exhibit a form of consciousness. If that seems a little far-fetched, we'll soon know for sure. We're about to start building artificial brains, and they stand an even higher chance of lighting up with some kind of conscious thought. Before we explore that frontier, though, we need to consider a rather provocative question.

'Can Machines Think?' It's a good question, which is why Hugh Loebner put it on his gold medal. Loebner is quite a character, by all accounts. He is a tireless champion of the rights of sex workers and, if you believe the press, a control freak obsessed with achieving Artificial Intelligence (AI). It's ironic, then, that the annual Loebner Prize is a competition that has serious AI researchers frothing at the mouth.

The premise is simple. Loebner will give $100,000 to the creator of a computer that can hold a typed conversation that is indistinguishable from conversation with a human. The prize comes with an 18-carat solid gold medal, on one side of which the face of Alan Turing appears, accompanied by Turing's provocative question.

The question is taken directly from a paper Turing published in 1950, 'Computing Machinery and Intelligence'. Turing outlined the idea that a sufficiently proficient machine, appropriately programmed, might be able to hold a conversation – via teleprinter – in a way that made it seem human. He called his suggestion the 'Imitation Game' and proposed it as a thought experiment (initially, at least) to address the question of whether the human brain does things a machine can't.

It is hard to know whether the Turing Test is of real value, but the annual jamboree surrounding Loebner's competition, which first ran in 1991, has certainly exposed the paucity of our achievements in AI. No one has come close to winning the gold medal, and the conversations conducted by each year's best entrant (winning a bronze medal and $2,000) are almost farcically odd. Loebner's insistence on carrying on with the competition, despite the unimpressive results, has caused AI pioneer Marvin Minsky to offer $100 to anyone who can persuade Loebner to put it out of its misery. Loebner, though, is having none of it. (He is a difficult man to deal with, by all accounts. An article in *Salon* from 2003 puts it beautifully: 'By tradition, three things happen at the conclusion of every Loebner contest: The winners take their prizes and run for the nearest exit, Hugh Loebner basks in glory, and the hosting organization takes a solemn oath: "Never again.'")

The test wasn't Turing's first foray into machine intelligence. In 1948, while still based at the National Physical Laboratory in London, he wrote a fascinating paper he called 'Intelligent Machinery'. It describes an 'unorganised machine', where artificial neurons are linked by random connections that can be modified as necessary. He showed that, if the network was sufficiently large, it could perform all the functions of a general-purpose computer. Here, Turing suggested, was a possible link between the messy human brain and the high-level processing it manages to carry out.

It was not well received by his boss, Sir Charles Darwin (grandson of the famous Charles Darwin), who called it a 'schoolboy's essay' and left it to languish, unpublished, until fourteen years after Turing's death. The trouble was, computing machines were simply not advanced enough for anyone to take the possibility of mimicking human intelligence seriously.

In many ways the Turing Test misses the interesting point about machines that can think – perhaps deliberately. Biographer Andrew Hodges suggests that the Turing Test concept is a dodge that 'allows Turing to avoid any discussion of what consciousness is.' Perhaps to dabble in machine consciousness would have been a step too far for someone who was already seen as a dangerous maverick. After all, that 1948 paper does begin with what must have sounded like a ridiculous proposition.

big – neurons are only 20 micrometres in diameter – but that's plenty big enough when you can map things one thousand times smaller. We are gradually pulling together the information needed to create the neuronal map of a mouse brain – that's roughly 75 million neurons. The human brain's 86 billion neurons is a long way beyond that, but not so far that we can't set our sights on it.

The ambition is breathtaking. It will be possible, with advancing technology and the billion Euros of funding that has already been secured and started to flow into research labs in Europe, to simulate the human brain within ten years. Computing power doubles every eighteen months, and the massive computing power required for the Human Brain Project – roughly a thousand times what is available now – should be online by 2023. That will be enough to model the detail of neuronal connections, and the way ions flow in and out to pass signals between both single neurons and different areas of the brain.

The stated goal of the Human Brain Project is to investigate what happens when a brain goes wrong, in afflictions such as Alzheimer's disease and Parkinson's disease. The unspoken promise is that a silicon brain may well show a form of consciousness. Which brings us back to the zombie.

We used to think consciousness was all about the cortex, that this recent evolutionary addition to the brain was the root of consciousness. However, we now know that creatures without a cortex make conscious decisions and display emotional states. As Mark Bekoff and Jessica Pierce observed in their book, *Animal Justice*, 'We are not the only moral beings.' The Cambridge Declaration on Consciousness makes that crystal clear: 'The neural substrates of emotions do not appear to be confined to cortical structures… Wherever in the brain one evokes instinctual emotional behaviors in non-human animals, many of the ensuing behaviors are consistent with experienced feeling states, including those internal states that are rewarding and punishing. Deep brain stimulation of these systems in humans can also generate similar affective states.' Consciousness happens in the brain – all over the brain. Which means that our built brains might show it arising, killing that zombie dead.

As Stanford University philosopher Paul Skokowski demonstrates in his essay, *I, Zombie*, 'Microphysical and functional duplicates of us living in a duplicate world will have conscious experience just like the beings that inhabit this world: *us*.' This means, he points out, that zombies – duplicates without any conscious experience whatsoever – are impossible. 'So please,' Skokowski adds, 'do not bother weeping for your zombie duplicate; after all, he or she will be feeling the same pain as you.' It's an elegant rephrasing of what Patricia Churchland calls, slightly less elegantly, the 'hornswoggle problem'. Chalmers' zombie argument, she says, is 'hornswoggle': bamboozling, sleight of hand, no different to the tricks that Gustav Kuhn performs. Why? Because we might well solve the easy problems of consciousness: how we form memories that we can dredge up into our consciousness; how our visual perception becomes a conscious observation, a noted sighting; the difference between unconscious sleeping states and the experience of being awake; how we pay attention to something; how pain is sometimes only experienced once we have had our attention drawn to what is causing it – and then find there is no 'Hard Problem' left to solve.

Churchland compares the objections people have had to this notion with the objections people had to other revolutionary ideas. 'When people were told that the earth moves they thought this was hilarious; it was ludicrous; it was inconceivable,' she says. Perhaps the best analogy is the problem people had seeing light as an electromagnetic wave. Light had religious and emotional significance, and the idea that it came from the same phenomenon that caused magnetism or static electricity was almost demeaning. The idea that our consciousness is a product of electrical interactions between neurons is similarly problematic. But we'll get used to it. In just a few decades we'll marvel at the problem people had with consciousness, she reckons.

And the zombie will be dead. Interpretations of the psychology associated with brains, coupled with the increased understanding of what goes on at the cellular or even molecular level, will give us the insights equivalent to our modern understanding of the partnership between electromagnetic waves and the phenomenon we know as light. The neuronal firings and the consciousness are inseparable; the zombie, the collection of neurons without the consciousness, simply vanishes in a puff of logic. 'They're not two things embracing each other, they're actually just one thing, looked at from two different points of view,' is how Patricia's husband, neuroscientist and philosopher Paul Churchland, puts it.

It is time to let go of the rigid definitions and embrace the complexity. Our over-confidence in interpreting outward signs or simplistic criteria have led us down a blind alley with our investigations of consciousness. We became convinced that there must be some secret self, some hidden extra for true consciousness to exist. But in the end, all you seem to need is some kind of information processing equipment that works in a similar way to the jelly-like mass inside our skulls. If you've got what we'd call a brain – whatever form it takes – you'll exhibit some kind of consciousness. Not all of the brain needs to work for conscious activity. In some cases, relatively little of it needs to work. What's more, different-sized brains are likely to exhibit different kinds of consciousness, but we aren't the arbiters of the cut-off point where an organism isn't conscious. Blue light can't tell red light it isn't light: it's a spectrum, a spread. And to think, we are about to build ourselves a brain, perhaps creating a new consciousness somewhere on that spectrum, a consciousness the like of which has never been seen on Earth. Could there be a more exciting moment?

To move us on, here is an observation that neatly sums up the dilemma of neuroscientists wanting to understand what makes us the people we are: 'Currently, invasive procedures, such as decapitation, are required for most genomic analyses of brain processes, ruling out human studies.' Yes, it would be a boon to know what's going on in there, but decapitating people to find out is probably going too far. But is decapitating animals any better?

The quote comes from an article by Samuel D. Gosling and Pranjal H. Mehta in a book called *Animal Personalities*. Even the researchers working in this field aren't always comfortable