

The Mystery of Consciousness

By Steven Pinker

The young woman had survived the car crash, after a fashion. In the five months since parts of her brain had been crushed, she could open her eyes but didn't respond to sights, sounds or jabs. In the jargon of neurology, she was judged to be in a persistent vegetative state. In crueller everyday language, she was a vegetable.

So picture the astonishment of British and Belgian scientists as they scanned her brain using a kind of MRI that detects blood flow to active parts of the brain. When they recited sentences, the parts involved in language lit up. When they asked her to imagine visiting the rooms of her house, the parts involved in navigating space and recognizing places ramped up. And when they asked her to imagine playing tennis, the regions that trigger motion joined in. Indeed, her scans were barely different from those of healthy volunteers. The woman, it appears, had glimmerings of consciousness.

Try to comprehend what it is like to be that woman. Do you appreciate the words and caresses of your distraught family while racked with frustration at your inability to reassure them that they are getting through? Or do you drift in a haze, springing to life with a concrete thought when a voice prods you, only to slip back into blankness? If we could experience this existence, would we prefer it to death? And if these questions have answers, would they change our policies toward unresponsive patients--making the Terri Schiavo case look like child's play?

The report of this unusual case last September was just the latest shock from a bracing new field, the science of consciousness. Questions once confined to theological speculations and late-night dorm-room bull sessions are now at the forefront of cognitive neuroscience. With some problems, a modicum of consensus has taken shape. With others, the puzzlement is so deep that they may never be resolved. Some of our deepest convictions about what it means to be human have been shaken.

It shouldn't be surprising that research on consciousness is alternately exhilarating and disturbing. No other topic is like it. As René Descartes noted, our own consciousness is the most indubitable thing there is. The major religions locate it in a soul that survives the body's death to receive its just deserts or to meld into a global mind. For each of us, consciousness is life itself, the reason Woody Allen said, "I don't want to achieve immortality through my work. I want to achieve it by not dying." And the conviction that other people can suffer and flourish as each of us does is the essence of empathy and the foundation of morality.

To make scientific headway in a topic as tangled as consciousness, it helps to clear away some red herrings. Consciousness surely does not depend on language. Babies, many animals and patients robbed of speech by brain damage are not insensate robots; they have reactions like ours that indicate

that someone's home. Nor can consciousness be equated with self-awareness. At times we have all lost ourselves in music, exercise or sensual pleasure, but that is different from being knocked out cold.

THE "EASY" AND "HARD" PROBLEMS

What remains is not one problem about consciousness but two, which the philosopher David Chalmers has dubbed the Easy Problem and the Hard Problem. Calling the first one easy is an in-joke: it is easy in the sense that curing cancer or sending someone to Mars is easy. That is, scientists more or less know what to look for, and with enough brainpower and funding, they would probably crack it in this century.

What exactly is the Easy Problem? It's the one that Freud made famous, the difference between conscious and unconscious thoughts. Some kinds of information in the brain--such as the surfaces in front of you, your daydreams, your plans for the day, your pleasures and peeves--are conscious. You can ponder them, discuss them and let them guide your behavior. Other kinds, like the control of your heart rate, the rules that order the words as you speak and the sequence of muscle contractions that allow you to hold a pencil, are unconscious. They must be in the brain somewhere because you couldn't walk and talk and see without them, but they are sealed off from your planning and reasoning circuits, and you can't say a thing about them.

The Easy Problem, then, is to distinguish conscious from unconscious mental computation, identify its correlates in the brain and explain why it evolved.

The Hard Problem, on the other hand, is why it feels like something to have a conscious process going on in one's head--why there is first-person, subjective experience. Not only does a green thing look different from a red thing, remind us of other green things and inspire us to say, "That's green" (the Easy Problem), but it also actually looks green: it produces an experience of sheer greenness that isn't reducible to anything else. As Louis Armstrong said in response to a request to define jazz, "When you got to ask what it is, you never get to know."

The Hard Problem is explaining how subjective experience arises from neural computation. The problem is hard because no one knows what a solution might look like or even whether it is a genuine scientific problem in the first place. And not surprisingly, everyone agrees that the hard problem (if it is a problem) remains a mystery.

Although neither problem has been solved, neuroscientists agree on many features of both of them, and the feature they find least controversial is the one that many people outside the field find the most shocking. Francis Crick called it "the astonishing hypothesis"--the idea that our thoughts, sensations,

joys and aches consist entirely of physiological activity in the tissues of the brain. Consciousness does not reside in an ethereal soul that uses the brain like a PDA; consciousness is the activity of the brain.

THE BRAIN AS MACHINE

Scientists have exorcised the ghost from the machine not because they are mechanistic killjoys but because they have amassed evidence that every aspect of consciousness can be tied to the brain. Using functional MRI, cognitive neuroscientists can almost read people's thoughts from the blood flow in their brains. They can tell, for instance, whether a person is thinking about a face or a place or whether a picture the person is looking at is of a bottle or a shoe.

And consciousness can be pushed around by physical manipulations. Electrical stimulation of the brain during surgery can cause a person to have hallucinations that are indistinguishable from reality, such as a song playing in the room or a childhood birthday party. Chemicals that affect the brain, from caffeine and alcohol to Prozac and LSD, can profoundly alter how people think, feel and see. Surgery that severs the corpus callosum, separating the two hemispheres (a treatment for epilepsy), spawns two consciousnesses within the same skull, as if the soul could be cleaved in two with a knife.

And when the physiological activity of the brain ceases, as far as anyone can tell the person's consciousness goes out of existence. Attempts to contact the souls of the dead (a pursuit of serious scientists a century ago) turned up only cheap magic tricks, and near death experiences are not the eyewitness reports of a soul parting company from the body but symptoms of oxygen starvation in the eyes and brain. In September, a team of Swiss neuroscientists reported that they could turn out-of-body experiences on and off by stimulating the part of the brain in which vision and bodily sensations converge.

THE ILLUSION OF CONTROL

Another startling conclusion from the science of consciousness is that the intuitive feeling we have that there's an executive "I" that sits in a control room of our brain, scanning the screens of the senses and pushing the buttons of the muscles, is an illusion. Consciousness turns out to consist of a maelstrom of events distributed across the brain. These events compete for attention, and as one process outshouts the others, the brain rationalizes the outcome after the fact and concocts the impression that a single self was in charge all along.

Take the famous cognitive-dissonance experiments. When an experimenter got people to endure electric shocks in a sham experiment on learning, those who were given a good rationale ("It will help scientists understand learning") rated the shocks as more painful than the ones given a feeble rationale ("We're curious.") Presumably, it's because the second group would have felt foolish to have suffered

for no good reason. Yet when these people were asked why they agreed to be shocked, they offered bogus reasons of their own in all sincerity, like "I used to mess around with radios and got used to electric shocks."

It's not only decisions in sketchy circumstances that get rationalized but also the texture of our immediate experience. We all feel we are conscious of a rich and detailed world in front of our eyes. Yet outside the dead center of our gaze, vision is amazingly coarse. Just try holding your hand a few inches from your line of sight and counting your fingers. And if someone removed and reinserted an object every time you blinked (which experimenters can simulate by flashing two pictures in rapid sequence), you would be hard pressed to notice the change. Ordinarily, our eyes flit from place to place, alighting on whichever object needs our attention on a need-to-know basis. This fools us into thinking that wall-to-wall detail was there all along--an example of how we overestimate the scope and power of our own consciousness.

Our authorship of voluntary actions can also be an illusion, the result of noticing a correlation between what we decide and how our bodies move. The psychologist Dan Wegner studied the party game in which a subject is seated in front of a mirror while someone behind him extends his arms under the subject's armpits and moves his arms around, making it look as if the subject is moving his own arms. If the subject hears a tape telling the person behind him how to move (wave, touch the subject's nose and so on), he feels as if he is actually in command of the arms.

The brain's spin doctoring is displayed even more dramatically in neurological conditions in which the healthy parts of the brain explain away the foibles of the damaged parts (which are invisible to the self because they are part of the self). A patient who fails to experience a visceral click of recognition when he sees his wife but who acknowledges that she looks and acts just like her deduces that she is an amazingly well-trained impostor. A patient who believes he is at home and is shown the hospital elevator says without missing a beat, "You wouldn't believe what it cost us to have that installed."

Why does consciousness exist at all, at least in the Easy Problem sense in which some kinds of information are accessible and others hidden? One reason is information overload. Just as a person can be overwhelmed today by the gusher of data coming in from electronic media, decision circuits inside the brain would be swamped if every curlicue and muscle twitch that was registered somewhere in the brain were constantly being delivered to them. Instead, our working memory and spotlight of attention receive executive summaries of the events and states that are most relevant to updating an understanding of the world and figuring out what to do next. The cognitive psychologist Bernard Baars likens consciousness to a global blackboard on which brain processes post their results and monitor the results of the others.

BELIEVING OUR OWN LIES

A second reason that information may be sealed off from consciousness is strategic. Evolutionary biologist Robert Trivers has noted that people have a motive to sell themselves as beneficent, rational, competent agents. The best propagandist is the one who believes his own lies, ensuring that he can't leak his deceit through nervous twitches or self-contradictions. So the brain might have been shaped to keep compromising data away from the conscious processes that govern our interaction with other people. At the same time, it keeps the data around in unconscious processes to prevent the person from getting too far out of touch with reality.

What about the brain itself? You might wonder how scientists could even begin to find the seat of awareness in the cacophony of a hundred billion jabbering neurons. The trick is to see what parts of the brain change when a person's consciousness flips from one experience to another. In one technique, called binocular rivalry, vertical stripes are presented to the left eye, horizontal stripes to the right. The eyes compete for consciousness, and the person sees vertical stripes for a few seconds, then horizontal stripes, and so on.

A low-tech way to experience the effect yourself is to look through a paper tube at a white wall with your right eye and hold your left hand in front of your left eye. After a few seconds, a white hole in your hand should appear, then disappear, then reappear.

Monkeys experience binocular rivalry. They can learn to press a button every time their perception flips, while their brains are impaled with electrodes that record any change in activity. Neuroscientist Nikos Logothetis found that the earliest way stations for visual input in the back of the brain barely budged as the monkeys' consciousness flipped from one state to another. Instead, it was a region that sits further down the information stream and that registers coherent shapes and objects that tracks the monkeys' awareness. Now this doesn't mean that this place on the underside of the brain is the TV screen of consciousness. What it means, according to a theory by Crick and his collaborator Christof Koch, is that consciousness resides only in the "higher" parts of the brain that are connected to circuits for emotion and decision making, just what one would expect from the blackboard metaphor.

WAVES OF BRAIN

Consciousness in the brain can be tracked not just in space but also in time. Neuroscientists have long known that consciousness depends on certain frequencies of oscillation in the electroencephalograph (EEG). These brain waves consist of loops of activation between the cortex (the wrinkled surface of the brain) and the thalamus (the cluster of hubs at the center that serve as input-output relay stations). Large, slow, regular waves signal a coma, anesthesia or a dreamless sleep; smaller, faster, spikier ones correspond to being awake and alert. These waves are not like the useless hum from a noisy

appliance but may allow consciousness to do its job in the brain. They may bind the activity in far-flung regions (one for color, another for shape, a third for motion) into a coherent conscious experience, a bit like radio transmitters and receivers tuned to the same frequency. Sure enough, when two patterns compete for awareness in a binocular-rivalry display, the neurons representing the eye that is "winning" the competition oscillate in synchrony, while the ones representing the eye that is suppressed fall out of synch.

So neuroscientists are well on the way to identifying the neural correlates of consciousness, a part of the Easy Problem. But what about explaining how these events actually cause consciousness in the sense of inner experience--the Hard Problem?

TACKLING THE HARD PROBLEM

To appreciate the hardness of the hard problem, consider how you could ever know whether you see colors the same way that I do. Sure, you and I both call grass green, but perhaps you see grass as having the color that I would describe, if I were in your shoes, as purple. Or ponder whether there could be a true zombie--a being who acts just like you or me but in whom there is no self actually feeling anything. This was the crux of a Star Trek plot in which officials wanted to reverse-engineer Lieut. Commander Data, and a furious debate erupted as to whether this was merely dismantling a machine or snuffing out a sentient life.

No one knows what to do with the Hard Problem. Some people may see it as an opening to sneak the soul back in, but this just relabels the mystery of "consciousness" as the mystery of "the soul"--a word game that provides no insight.

Many philosophers, like Daniel Dennett, deny that the Hard Problem exists at all. Speculating about zombies and inverted colors is a waste of time, they say, because nothing could ever settle the issue one way or another. Anything you could do to understand consciousness--like finding out what wavelengths make people see green or how similar they say it is to blue, or what emotions they associate with it--boils down to information processing in the brain and thus gets sucked back into the Easy Problem, leaving nothing else to explain. Most people react to this argument with incredulity because it seems to deny the ultimate undeniable fact: our own experience.

The most popular attitude to the Hard Problem among neuroscientists is that it remains unsolved for now but will eventually succumb to research that chips away at the Easy Problem. Others are skeptical about this cheery optimism because none of the inroads into the Easy Problem brings a solution to the Hard Problem even a bit closer. Identifying awareness with brain physiology, they say, is a kind of "meat chauvinism" that would dogmatically deny consciousness to Lieut. Commander Data just because he doesn't have the soft tissue of a human brain. Identifying it with information processing

would go too far in the other direction and grant a simple consciousness to thermostats and calculators--a leap that most people find hard to stomach. Some mavericks, like the mathematician Roger Penrose, suggest the answer might someday be found in quantum mechanics. But to my ear, this amounts to the feeling that quantum mechanics sure is weird, and consciousness sure is weird, so maybe quantum mechanics can explain consciousness.

And then there is the theory put forward by philosopher Colin McGinn that our vertigo when pondering the Hard Problem is itself a quirk of our brains. The brain is a product of evolution, and just as animal brains have their limitations, we have ours. Our brains can't hold a hundred numbers in memory, can't visualize seven-dimensional space and perhaps can't intuitively grasp why neural information processing observed from the outside should give rise to subjective experience on the inside. This is where I place my bet, though I admit that the theory could be demolished when an unborn genius--a Darwin or Einstein of consciousness--comes up with a flabbergasting new idea that suddenly makes it all clear to us.

Whatever the solutions to the Easy and Hard problems turn out to be, few scientists doubt that they will locate consciousness in the activity of the brain. For many nonscientists, this is a terrifying prospect. Not only does it strangle the hope that we might survive the death of our bodies, but it also seems to undermine the notion that we are free agents responsible for our choices--not just in this lifetime but also in a life to come. In his millennial essay "Sorry, but Your Soul Just Died," Tom Wolfe worried that when science has killed the soul, "the lurid carnival that will ensue may make the phrase 'the total eclipse of all values' seem tame."

TOWARD A NEW MORALITY

My own view is that this is backward: the biology of consciousness offers a sounder basis for morality than the unprovable dogma of an immortal soul. It's not just that an understanding of the physiology of consciousness will reduce human suffering through new treatments for pain and depression. That understanding can also force us to recognize the interests of other beings--the core of morality.

As every student in Philosophy 101 learns, nothing can force me to believe that anyone except me is conscious. This power to deny that other people have feelings is not just an academic exercise but an all-too-common vice, as we see in the long history of human cruelty. Yet once we realize that our own consciousness is a product of our brains and that other people have brains like ours, a denial of other people's sentience becomes ludicrous. "Hath not a Jew eyes?" asked Shylock. Today the question is more pointed: Hath not a Jew--or an Arab, or an African, or a baby, or a dog--a cerebral cortex and a thalamus? The undeniable fact that we are all made of the same neural flesh makes it impossible to deny our common capacity to suffer.

And when you think about it, the doctrine of a life-to-come is not such an uplifting idea after all because it necessarily devalues life on earth. Just remember the most famous people in recent memory who acted in expectation of a reward in the hereafter: the conspirators who hijacked the airliners on 9/11.

Think, too, about why we sometimes remind ourselves that "life is short." It is an impetus to extend a gesture of affection to a loved one, to bury the hatchet in a pointless dispute, to use time productively rather than squander it. I would argue that nothing gives life more purpose than the realization that every moment of consciousness is a precious and fragile gift.

*Steven Pinker is Johnstone Professor of Psychology at Harvard and the author of *The Language Instinct*, *How the Mind Works* and *The Blank Slate**