What is consciousness? Science faces its hardest problem yet

Allison Parshall February 2026

Until half a billion years ago, life on Earth was slow. The seas were home to single-celled microbes and largely stationary soft-bodied creatures. But at the dawn of the Cambrian era, some 540 million years ago, everything exploded. Bodies diversified in all directions, and many organisms developed appendages that let them move quickly around their environment. These ecosystems became competitive places full of predators and prey. And our branch of the tree of life evolved an incredible structure to navigate it all: the brain.

We don't know whether this was the moment when consciousness first arose on Earth. But it might have been when living creatures began to really need something like it to combine a barrage of sensory information into one unified experience that could guide their actions. It's because of this ability to *experience* that, eventually, we began to feel pain and pleasure. Eventually, we became guided not just by base needs but by curiosity, emotions and introspection. Over time we became aware of ourselves.

This last step is what we have to thank for most of art, science and philosophy—and the millennia-long quest to understand consciousness itself. This state of awareness of ourselves and our environment comes with many mysteries. Why does being awake and alive, being *yourself*, feel like anything at all, and where does this singular sense of awareness come from in the brain? These questions may have objective answers, but because they are about private, subjective experiences that can't be directly measured, they exist at the very boundaries of what the scientific method can reveal.

Still, in the past 30 years neuroscientists scouring the brain for the so-called neural correlates of consciousness have learned a lot. Their search has revealed constellations of brain networks whose connections help to explain what happens when we lose consciousness. We now have troves of data and working theories, some with mind-bending implications. We have tools to help us detect consciousness in people with brain injuries. But we still don't have easy answers—researchers can't even agree on what consciousness is, let alone how best to reveal its secrets. The past few years have seen accusations of pseudoscience, results that challenge leading theories, and the uneasy feeling of a field at a crossroads.

Yet the stakes for understanding consciousness have never been higher. We've built talking machines able to imitate consciousness so well that we can't always tell the difference. Sometimes these artificial-intelligence models claim outright to be sentient. Faced with an existential unknown, the public is turning to the field of consciousness science for answers. "The tension, you know, it's palpable," says Marcello Massimini, a neurophysiologist at the University of Milan. "We're going to be looking back at this period."

Consciousness is all you really know. It's the voice you hear in your head, your emotions, your awareness of the world and your body all rolled into one unified experience. "Everything comes down to it, everything," says cognitive neuroscientist Athena Demertzi of the University of Liège in Belgium. "It's the translation of the world that we have." Philosophers and scientists alike struggle to define consciousness without appealing back to what it *feels* like to experience—what philosophers call "definition by pointing." But they're pointing to a real phenomenon. It's your consciousness that goes wonky when you take hallucinogens, even as your body and environment stay the same. When you go under general anesthesia, it appears to go out like a light. When you dream, some strange form of consciousness persists, even if it's disconnected from the outside world.

Some scientists have used these different states of consciousness to chop up conscious experience into at least three pieces: wakefulness, internal awareness and connectedness. In a “normal” state of consciousness, you have all three. You’re awake with your eyes open, a state that is sustained by signals from your brain stem. You’re internally aware, forming thoughts and mental imagery. And you’re connected to the outside world, with your brain receiving and processing information from the five senses.

Jen Christiansen; Source: “Near-Death Experience as a Probe to Explore (Disconnected) Consciousness,” by Charlotte Martial et al., in *Trends in Cognitive Sciences*, Vol. 24; March 2020 (*primary framework reference*)

How the brain gives rise to these strange experiences is a question that has haunted neuroscience for as long as the field has existed. Massimini was driven wild by the mystery in medical school, when he held a brain in his hands for the first time. “This is an object with boundaries, with a given weight, a little bit like tofu. It’s not particularly elegant,” he says, but “inside this object that you can hold in your hand, there is a universe.” Many philosophical traditions have dealt with this apparent disconnect by saying the mind—or the soul—is not made of the same physical stuff as our bodies, a position called dualism. Science has instead flourished by assuming the opposite and siding with a theory called materialism, which presumes that everything we observe somehow arises from physical matter, including consciousness.

Perhaps knowing they weren’t up to the job of explaining how this happens, neuroscientists shied away from the enigmas of consciousness until the 1990s. “You had to be retired or religious or a philosopher to be able to talk about it,” says neuroscientist Christof Koch, a member of the *Scientific American* board of advisers and chief scientist at the Tiny Blue Dot Foundation, a research nonprofit focusing on perception science. In 1990 Koch and Nobel laureate Francis Crick, the co-discoverer of DNA, directly challenged this taboo. They published a paper that laid out an ambitious plan to study the neurobiology of consciousness and launched the field as we know it today.

Their plan came at a good time. That same year neuroscientists invented a new way to observe the working brain called functional magnetic resonance imaging (fMRI). Using brain scanners, they track changes in blood flow to reveal which brain regions are active at a given time, producing colorful images of the brain in action. Koch, who studied vision, thought that by measuring people's brain responses as they looked at special optical illusions, scientists could figure out which parts of the brain are activated when something is consciously perceived. Some of the illusions used can be perceived in one of two ways; one example is Rubin's vase, which can be seen either as a vase or as two faces in profile. The image never changes, so the brain is always receiving the same information, but people's conscious experience of it can easily switch back and forth. Another visual test, called binocular rivalry, has a similar effect: each eye is shown a different image, and people perceive either one or the other but never a mix of the two. If neuroscientists could scan people's brains as their conscious perception switched, they could find parts of the brain that were associated with that change: the neural correlates of consciousness.

Koch bet big, literally. In 1998, at a consciousness-science conference in Germany, he bet philosopher David Chalmers a case of wine that researchers would discover a "clear" pattern of brain activation underlying consciousness within

25 years. Chalmers took the bet, thinking 25 years "might be a bit optimistic," he says.

It was extremely optimistic. These early neural-correlate studies of vision, which dominated the field in the 1990s, suggested places that may be less correlated with visual awareness: the regions where input from our eyes first enters the brain. These low-level sensory-processing areas contain a lot of information that our conscious selves aren't privy to. These areas appear to continue receiving sensory information when we're under anesthesia as well. As that information travels "up" the wrinkly outer layer of the brain, called the cortex, it enters areas that pick out and process higher-level details—such as the faces in an image—and conscious awareness builds.

So consciousness seems to happen in some region outside these early visual-processing areas—but there is no consensus on where.

Today there are dozens of competing theories of how the brain generates consciousness. They have different starting points, different goals and even different definitions of consciousness. The most popular is global neuronal workspace theory (GNWT), which envisions consciousness as a kind of stage. When something enters your conscious awareness—an itch, say, or the buzz of your refrigerator—it's thrust onto the stage and spotlighted in a process called ignition. Things on the stage, or in the "global workspace," get broadcast to the rest of the brain, where they're able to guide action, direct attention, and more.

Higher-order theories conceive of consciousness as a high-level representation of what is going on in other parts of the brain. For you to be conscious of a refrigerator's buzzing, your brain cannot just represent the buzz by activating its auditory parts (which are located near the temples). The buzz must also have a corresponding "meta-representation" in the frontal parts of the brain that are responsible for higher-order thinking—such as the thought "I am hearing the refrigerator buzzing."

Reentry and predictive processing theories (PPTs), on the other hand, propose that consciousness emerges from our brain's balancing of two processes: perception and prediction. If you have ever seen something that wasn't there simply because you expected to see it, you know how much our brain's predictions can govern what we actually perceive. Neuroscientist Anil Seth of the University of Sussex in England, who favors PPTs, describes conscious perception as a "controlled hallucination," with the brain's best guess of what's going on around you mapping onto what you consciously perceive.

> *There's a chasm between our everyday experiences and what science can explain.*

Then there's integrated information theory (IIT), a mathematical and philosophical theory that stands out from the rest because it doesn't start with the brain. Instead it starts with consciousness itself and the observations we can make about its properties, then asks what kind of system could allow something with these properties to exist. IIT takes consciousness to be differentiated—there are a lot of things you could be experiencing right now but aren't, making your consciousness rich in information. And it is also unified, or integrated—all your diverse experiences are bundled into one single stream of consciousness. Mathematically, these two features together make the system very complex. And from this complexity comes consciousness.

Most eye-catching of all, IIT implies that consciousness could be present outside of living systems, a kind of panpsychism. This idea, plus the theory's relative lack of grounding in the brain and coverage in the media, would make IIT a flashpan of controversy. But first it would inspire one of the most important insights we have into how consciousness works.

In the early 2000s, while studying in the U.S., Massimini began performing experiments with a device for probing the brain that did two things at once: deliver painless magnetic pulses to the brain and detect brain waves, both from outside the scalp, techniques called transcranial magnetic stimulation (TMS) and electroencephalography (EEG), respectively. Once back in Italy, he managed to secure a grant to buy a TMS-EEG machine for his university despite a "desperate" research-funding situation in the country.

A few years later he and a colleague "did something crazy," he says. They loaded the machine into a truck and drove more than nine hours to Liège. "We didn't say anything to anybody. This is a machine belonging to the university after all." But the opportunity was too good to pass up. A neurologist in Liège named Steven Laureys had founded the Coma Science Group for treating and learning from patients with disorders of consciousness, and Massimini believed his new device could be used to measure someone's level of consciousness from their brain activity.

Researchers had tried to measure the difference between conscious and unconscious brains with other brain-imaging techniques before, without much success. But adding TMS let scientists stimulate the outer layers of the cortex, causing neurons in a specific area to fire. Then EEG measured brain waves to reveal how that stimulation spread. "It's like knocking on the brain directly," Massimini says, "to probe the internal structure."

You can also think of TMS like dropping a rock in a pond. In a conscious brain (whether awake or dreaming), the disturbance ripples outward as neurons cause neighbors in their networks to fire. But unlike waves in water, each of those ripples of neuron activity begets more ripples, spreading in a complex and far-reaching way throughout the brain's networks. In dreamless sleep, this doesn't happen, Massimini had previously found. TMS stimulates the brain, and the neurons fire, but the wave of activity isn't picked up by neighboring neurons. If there are ripples, they don't spread far. The complexity seen during wakefulness is gone.

In Liège, Massimini and his colleagues tested the technique on people with various disorders of consciousness—patients who were in vegetative states, or were in minimally conscious states, or were outwardly unresponsive but internally aware. They found that people whose brains exhibited a more complex response were more likely to be conscious. This relation could be represented as a single number, called the perturbational complexity index, or PCI.

PCI is a very crude measure of consciousness, but it can estimate someone's place on the spectrum of consciousness quite reliably. And it suggests that complexity is an important part of a conscious brain. In an awake or dreaming brain, diverse networks of neurons are in constant back-and-forth communication with one another. In this way, conscious brain activity is both differentiated (or rich in information) and integrated (forming one unified whole)—principles that Massimini borrowed from IIT, the theory that doesn't begin with the brain. These interactions build up complexity, or what IIT theorists call a "cause-effect structure," so that when you stimulate one part of a conscious brain, other parts respond.

But during dreamless sleep or when someone is under anesthesia, all that communication goes away. "Everything collapses," Massimini says. "The cathedral falls apart." Slow brain waves travel across the cortex as neurons cycle rhythmically between two electric states. In the "silent periods" between the waves, neurons enter what's called a down state, in which they can't respond to electric signals from their neighbors. This state is why there's silence when you stimulate an unconscious brain with TMS: "No feedback, no unity, no complexity," he says.

Of course, this loss of complexity during sleep and anesthesia is transient; disorders of consciousness can be permanent. "Why can I reverse sleep in a few seconds, and I can reverse anesthesia in the course of minutes, but I might not ever be able to reverse this pathological state?" asks George Mashour, an anesthesiologist and neuroscientist studying consciousness at the University of Michigan Medical School. Massimini hopes that we can eventually learn how to jump-start consciousness—rebuild the cathedral—for people who are in vegetative or minimally conscious states.

> *"It feels like there's been a hard-won legitimacy to the study of consciousness over the past 30 years." —Anil Seth, University of Sussex*

Yet understanding brain-network complexity does not solve the mystery of consciousness. These findings can help explain how a brain can reach the state of consciousness but not what happens once it's gotten there, Mashour points out. Changes in someone's PCI value can't explain, for example, why The Dress looks blue and black one moment and white and gold the next. It can't explain how a toothache feels different from a headache, how someone without functioning circulation can have a near-death experience, or how the psychedelic drug 5-MeO-DMT makes time seem to stop and obliterates your sense of self.

There's a chasm between our everyday experiences and what science can explain. "No one really has a theory that closes the explanatory gap," says Tim Bayne, a philosopher at Monash University in Melbourne. "But that's a problem on us, not the brain."

At a June 2023 conference in New York City, Koch gave Chalmers his case of wine and conceded that he had lost their bet. "It's clear that things are not clear," Chalmers said.

That weekend the evidence looked particularly murky. The results of a massive research project pitting IIT against GNWT had recently been shared. The project, led by a group called the Cogitate Consortium, involved three different measuring techniques used in eight different institutions around the world. Researchers developed predictions from each theory about what should happen in the brain when an image is consciously perceived versus when it's not. Testing those predictions could challenge or even falsify either theory.

Both theories came away bruised. IIT holds that consciousness arises mostly from sustained activity in the back of the brain. This "hot zone" sits at the intersection of many sensory networks of neurons. GNWT, in contrast, predicts that a stimulus (such as an image) rises to the level of consciousness only when there is an "ignition" to the workspace in frontal parts of the brain such as the prefrontal cortex, which is known for planning and decision-making. GNWT also predicts that this ignition signal will appear as two discrete spikes of activity—one when an image is first presented and one when it's removed—whereas IIT predicts sustained activity as long as a person is looking at an image.

The results were extremely mixed. Although there was sustained activity in the back of the brain associated with conscious perception, networks in the region weren't synchronized in the way that IIT predicts. And although there was a signal in the prefrontal cortex when images were first presented, there was not a second signal when they were removed, contrary to GNWT's predictions.

Then, a few months later, the field erupted. An open letter calling IIT pseudoscience was published online in September 2023, signed by 124 researchers in or adjacent to the field. The argument focused less on the theory than on its coverage in the media, which the letter's authors saw as credulous. The authors also took issue with the panpsychist implications of IIT, highlighting descriptions of it as unscientific and "magicalist." "These bold claims threaten to delegitimize the scientific study of consciousness," many of the authors wrote in a follow-up article.

The prospect that the field could lose its legitimacy hung over the fight. One side feared IIT's reputation would drag consciousness science even further toward the fringes, and the other worried that publicly tarring one theory with a "pseudoscience" label would lead to the downfall of the entire field. “My greatest fear is that we get another ‘consciousness winter’ wherein just talking about consciousness is considered pseudoscientific bunk,” wrote Erik Hoel, a consciousness researcher at Tufts University who has published extensively on IIT’s limitations, in a post defending the theory.

The debate, which took place largely in online posts and in the media, was finally hashed out in the pages of *Nature Neuroscience* last March. Since then, the scientists involved have seemed to be trying to put the ugly chapter behind them. But now there is a sense that the field has arrived at an “uneasy stasis,” Seth and his colleagues wrote recently in the journal *Frontiers in Science*.

“It feels like there’s been a hard-won legitimacy to the study of consciousness over the past 30 years,” Seth says. And there are important results to show for it. We now know that large parts of the brain—for example, the cerebellum, a structure near the brain stem that contains a majority of the brain’s neurons—is apparently not involved in consciousness. We’ve learned about specific brain regions that are associated with specific pieces of conscious experience, such as our sense of self. We’re also getting hints that ancient structures deep in our brain, such as the thalamus, may be more involved than neuroscientists had previously thought.

Comparing consciousness among species could reveal why it exists in the first place.

But underneath it all lurk countless unknowns. “There’s still disagreement about how to define [consciousness], whether it exists or not, whether a science of consciousness is really possible or not, whether we’ll be able to say anything about consciousness in unusual situations like [artificial intelligence],” Seth says. It stands in contrast, perhaps unfairly, to other scientific journeys of discovery, such as the mapping of our genetic code in the Human Genome Project or of the cosmos with the help of the James Webb Space Telescope, he adds.

“It’s a wonderful moment but also kind of sobering,” Bayne says. Building bigger and bigger particle colliders is a pretty good tactic for revealing the stuff of the subatomic world. But for revealing the stuff of consciousness, there’s no sure bet. “If Bill Gates gave me $100 billion tomorrow and said, ‘Okay, find out about consciousness,’” he says, “I wouldn’t know what to do with that money.”

Artificial intelligence may soon force our hand. In 2022, when a Google engineer publicly claimed the AI model called LaMDA he had been developing appeared to be conscious, Google countered that there was “no evidence that LaMDA was sentient (and lots of evidence against it).” This struck Chalmers as odd: What evidence could the company have been talking about? “No one can say for sure they’ve demonstrated these systems are not conscious,” he says. “We don’t have that kind of proof.”

As these machines get better at imitating human dialogue—sometimes even claiming outright to be conscious—ethicists, AI companies and the concerned public are increasingly looking to consciousness research for answers. “Suddenly those philosophical questions have become very practical questions,” Chalmers says.

These questions are bigger and older than AI. Where does consciousness exist in the world around us, and how can we prove it? Scientists and philosophers are increasingly studying animals, human fetuses, brain organoids and AI to figure out what common principles could underlie consciousness.

Researchers have often studied consciousness by focusing exclusively on humans, because the only consciousness we can ever truly be sure exists is our own. For everyone else, we must rely on behavioral cues and trust they are not a “philosophical zombie,” with all the outward signs of consciousness but without any of the internal experience. We extend this assumption to fellow humans every day.

Sometime in the 20th century, though, scientists stopped doing so for animals. “When I started my graduate studies in the 1990s, ‘chimpanzees aren’t conscious’ was the default position for a lot of philosophers,” says Kristin Andrews, a philosopher studying animal minds at the City University of New York Graduate Center.

Jen Christiansen; Source: "Theories of Consciousness," by Anil K. Seth and Tim Bayne, in *Nature Reviews Neuroscience*, Vol. 23; July 2022 (*primary reference*)

Yet we find consciousness only where we presume to look for it. It's a spotlight effect, Andrews explains, and since then, our spotlight has slowly widened. First, in the 1990s, consciousness scientists broadened it to do research on lab monkeys that couldn't be done on humans. By the time a group of scientists signed the Cambridge Declaration on Consciousness, in 2012, there was more acceptance of the idea that all mammals and some birds are probably sentient.

Now the frontier rests with fish, crustaceans and insects. Studies suggest that fish can recognize themselves in a mirror, bumblebees can play and crabs can weigh decisions based on conflicting priorities. The 2024 New York Declaration on Animal Consciousness, which Andrews co-authored, states that there is at least a "realistic possibility" of consciousness in all vertebrates and some invertebrates, such as insects, certain mollusks and crustaceans. "We can't just assume that all these animals are not conscious," says Chalmers, who signed the declaration.

Comparing consciousness among species could reveal why it exists in the first place. "People have focused a lot on where consciousness is in the brain and perhaps less so on what it's *for*," Seth says. He theorizes that consciousness is intrinsically linked to life. Living beings can do only one thing at a time, and to choose what to do, they must bring a lot of relevant information together into one stream.

Even if that is right, it doesn’t mean carbon-based life is the only arena where consciousness can happen. “Just as we can build things that fly without flapping their wings, maybe there are other ways of being conscious that don’t require being alive,” Seth says. “We should really take that possibility seriously.”

The AI large language models (LLMs) that underpin chatbots such as ChatGPT and Claude can certainly imitate consciousness well, although today they are most likely the zombies that Chalmers and other philosophers once imagined. Even most AI enthusiasts will tell you that all an LLM does is predict which word comes next in a sentence; it doesn’t “know” anything. But—to be strictly philosophical about it—can we really *prove* LLMs aren’t conscious if we haven’t yet agreed on how consciousness works?

Some researchers think theories rooted in the human brain, such as GNWT, could still provide clues. If the brain is like a biological computer—a dominant assumption of cognitive neuroscience—then maybe researchers can compare how LLMs process information and test for indicators of consciousness. GNWT, which was itself inspired by an early type of AI model, says information is consciously experienced once it’s broadcast across the entire system.

Does an LLM do something similar?

Not everyone buys the brain-computer circuitry analogy. Brains do a lot more than run algorithms that process information, Seth says. They have electric fields, and they interact with chemical signals. They are made of thousands of types of living cells that consume energy. “It’s a massive assumption that none of these things matter,” he says. “And that assumption has gone largely unexamined because of the power of the metaphor that the brain is a computer.”

IIT proponents such as Massimini and Koch also think the underlying physical “stuff” of a system matters—and that mere simulations, including LLMs, can’t yield consciousness. “It’s like [how] simulating a storm will not get you wet,” Massimini says, “or simulating a black hole will not bend space and time.”

In consciousness science, everything comes back to the measurement problem. You can try to find markers of different states of consciousness—for instance, by scanning a person’s brain while they are awake versus in slow-wave sleep, which is typically dreamless and therefore unconscious. This experimental setup assumes the subject is in fact not dreaming. But that assumption could be wrong: sometimes people do report dreams when woken from slow-wave sleep. Were they wrong? Do you trust them? How can you confirm that your assumptions about consciousness are correct when your only ground truth is someone else's word —which is not really a ground truth at all?

When we are faced with this seemingly intractable problem, it's tempting to reach for an escape valve: Maybe none of it is real. Maybe consciousness is so illusory because it is an illusion, a beautiful cathedral that exists only in our heads. This skeptical position was often put forward by the late philosopher Daniel Dennett, and it's a legitimate question. But it doesn't allow us to opt out of treating brain injuries, understanding drugs such as anesthetics and psychedelics, and grappling with our treatment of animals and the intelligent machines we're birthing. Consciousness is real to us, and therefore it is real in every way that counts.

All of science rests on inferences about things we cannot see. We can't see a black hole, Koch points out, but we can spend decades building up theories and creating instruments that let us infer their existence. Consciousness may be a more challenging case, but researchers don't plan to stop trying. With the right tools, "the sense of mystery about how material processes could give rise to conscious experiences would start to go away," Seth says.

"I don't know what will happen afterward—if it will still be impressive or not," University of Liège's Demertzi says. "But, you know, sometimes nature is so beautiful that even when it's analyzed, you're in awe."