

How to think about consciousness

What is it like to be you? Dive into the philosophical puzzle of consciousness and see yourself and the world in new ways

By Amy Kind

Sep 04, 2024 10:00 AM · 12 min. read · [View original](#)

Think it through

Privacy is baked into consciousness

One unique aspect about conscious experience concerns the way that it is known. How do you know that a friend is undergoing emotional distress? By seeing the expressions on their face or listening to what they tell you. How do you know that you yourself are undergoing emotional distress? Here, you don't need to rely on any external cues. You have a kind of direct access to your own conscious states that you don't

have to your friend's conscious states. Your own conscious states are available to you by way of *introspection* – what we might metaphorically think of as a kind of 'looking within'.

The fact that we each have introspective access only to our own conscious states is intimately connected to a deep fact about the nature of consciousness: it is private. This privacy is a matter of principle, not of practice. That is, the privacy of conscious experience doesn't come down to a matter of personal choice. Rather, it is baked into the very essence of what consciousness is.

When you're sitting around a campfire with your friends, there's a sense in which your experience is shared with all of them – after all, you're seeing the same flames and hearing the same crackling logs. But there's another sense in which your conscious experience is not shared with anyone; it belongs only to you. When a friend tries to empathise with you, they might say: 'I feel your pain.' Of course, this shouldn't be taken literally. An individual cannot literally feel a pain that isn't theirs.

Or can they? We might put pressure on this idea by considering the case of Krista and Tatiana Hogan, the Canadian craniopagus twins who are fused at the skull. In everyday interactions, the Hogan girls reveal an incredible amount of mental interconnectedness. In fact, at times, Krista and Tatiana appear to be sharing their sensory experiences with one another. When they were babies, putting a pacifier in one twin's mouth could stop the other twin from crying, and one twin would show signs of feeling pain when the other was pricked by a needle for a blood draw. This connection has not seemed to lessen as they've grown. If their mother holds an object in front of one twin's eyes while the other's eyes are closed, the second twin can then report various facts about the object: what kind of toy animal it is, what its colour is, and so on. If one twin is touched on the leg or arm or face while the other twin's eyes are closed, the twin with closed eyes can report where her sister was being touched.

This fascinating case raises broader questions about whether technology may one day allow for some kind of mind

meld along the lines envisioned in the TV show *Star Trek* or for some other way for consciousness to be merged across different individuals. But, speculations about future technology aside, it remains true that as a general matter our conscious experiences are private to us.

Imagine being a bat

The essential privacy of conscious experience poses challenges for understanding others. The challenge is especially deep when it comes to others whose experiences are likely very different from our own. Have you ever been walking through a forest at night and caught a glimpse of a bat seamlessly navigating its way through the darkness? It seems like an impossibly alien thing to do. What would it be like to fly through the night like that? How could you ever figure out what it's like to be a bat? This is precisely the question asked by Thomas Nagel in a [paper](#) published 50 years ago.

Bats have conscious experience. They're mammals, and they engage in the kinds of sophisticated behaviour that we associate with consciousness. But bat

experience is very different from human experience. While humans navigate the world using sight and sound, bats navigate the world by way of echolocation. What it's like for a bat to employ echolocation is presumably vastly different from what it's like for a sighted person to employ vision. Might there be any way to close the gap between human experience and bat experience?

Given that humans don't echolocate (though perhaps some people use a similar [technique](#)), you can't have the same kinds of experience yourself that the bat has. Can you even *imagine* such experiences? Nagel thinks not:

It will not help to try to imagine that one has webbing on one's arms ... and perceives the surrounding world by a system of reflected high-frequency sound signals; and that one spends the day hanging upside down by one's feet in an attic. In so far as I can imagine this (which is not very far), it tells me only what it would be like for *me* to behave as a bat behaves. But that is not the question. I want to know what it is like for a *bat* to be a bat. Yet if I try to imagine this, I am restricted to the resources of my own mind, and those resources are inadequate to the task.

In Nagel's view, conscious experience is essentially subjective, essentially connected with only a single point of view. He also argues that it cannot be

explicated in objective terms. We will return to this latter point in a moment, but for now, let's note that it's the subjectivity of consciousness that makes it impossible for a person to achieve an understanding of conscious experiences that are vastly different from any that they've ever had.

When it comes to the experiences of people across vast experiential divides – people of a different race, ethnicity or gender, or of a different social class, or with a different ability status – knowing what it is like to be them might well be out of reach. Perhaps if you spend a lot of time listening to them, and you are a skilled imaginer, you might be able to leverage your imaginative capacities to achieve some understanding of their conscious experience, maybe even a high degree of understanding. But the subjectivity of consciousness makes this endeavour a very challenging one.

What colour do you see?

We all learn from an early age that, on the road, 'red' means stop and 'green' means go. When you see a red light, you put on the brakes, and when I see a red

light, I do too. And that leads us to believe that we are having conscious experiences that have the same phenomenal feeling, the same what-it's-like-ness. Philosophers often refer to these phenomenal features of experience as *qualia*. But here's a puzzling possibility: what if we are having very different colour qualia when we both look at the red light? Maybe when you look at a red light you are having the colour qualia that I have when I look at a green light, and vice versa. How could we ever tell one way or the other?

This puzzle was first raised back in the 17th century by the English philosopher John Locke, who noted that the very same object might produce different experiences in several people's minds at the same time without our being able to tell. As he explains the scenario, it could be that 'the idea that a violet produced in one man's mind by his eyes were the same that a marigold produced in another man's, and vice versa' and, moreover, 'this could never be known, because one man's mind could not pass into another man's body, to perceive

what appearances were produced by those organs.'

Philosophers refer to this disquieting possibility as *the inverted spectrum*, since what's proposed is that your qualia might be inverted in comparison with mine. The possibility of the inverted spectrum is closely related to the privacy and subjectivity of consciousness. It's precisely because conscious experiences can't be shared and can't be objectively captured that we cannot rule out the possibility that we have radically different colour qualia from one another, even when looking at the very same object.

It's tempting to think that our behaviour would reveal the difference. But reflect on how you learned colour terms. Your parent points to a ripe tomato or a stop sign or the muppet Elmo and tells you 'That's red.' They point to a stalk of broccoli or a grassy field or Kermit the frog and say 'That's green.' You naturally come to associate the word 'green' with the colour experience you are having when you see Kermit, however that colour experience feels to you – and even if the colour experience is the one that

your parent is having when they see Elmo. Whatever that colour experience is, you will develop into a flawless user of the words 'red' and 'green', you'll stop at red lights and go at green lights, and you'll correctly identify peas as being the same colour as broccoli, and apples as being the same colour as tomatoes.

Perhaps not all conscious experiences can be inverted without detection. For example, if we try to think about an inverted spectrum with respect to pleasure and pain, it's a lot harder to imagine how it could be undetectable. Suppose that when you're tickled by a feather you feel the kind of painful sensation that I feel when I step on a Lego piece. Since painful sensations are connected to involuntary bodily reactions like grimaces and wincing, it looks like this inversion would make itself apparent. But even if we can rule out inversion hypotheses for some types of conscious experiences, the puzzle of the inverted spectrum with respect to colour sensations still remains, and it leads to other questions about how different our conscious experiences might be from one another. It also poses challenges for

attempts to give scientific explanations of consciousness.

Science might not have all the answers

Consider a clock. You may not know exactly how it works, but if you were to take it apart and spend enough time thinking about it, you could probably come to understand quite a bit about its functioning. But now consider a human being. If somehow you were able to take it apart (and not in a weird, serial killer way), you might learn quite a bit about human functioning. But would you be able to learn what makes a human tick? Would you be able to learn anything about consciousness?

Many have thought the answer to this question is no. The subjectivity of consciousness not only poses a significant challenge to interpersonal understanding but it also poses a significant challenge to seeing how it fits in with the rest of the natural world. Other processes and activities in the natural world – like photosynthesis and erosion and precipitation – are wholly objective. Even other *human* processes and activities are wholly objective – like

digestion and reproduction and respiration. All of these objective processes can be fully explicated by science. Indeed, it's typically assumed that science will one day be able to explain *everything whatsoever*, it's also even assumed that science will one day provide us with a single, complete, unified theory of everything. But consciousness calls these assumptions into question. As Nagel's bat example or the inverted spectrum case seem to show, consciousness threatens to defy scientific explanation. Given how central and important consciousness is to our lives, this is an uncomfortable result.

On the one hand, it seems that consciousness clearly has something to do with the brain. When all brain activity ceases, as in brain death, conscious experience also ceases. Lesser forms of damage to the brain also have adverse effects on conscious experience. When individuals have lesions on one side of their primary visual cortex, for example, they lose visual consciousness of objects in that side of their visual field. Entities that don't have brains or any kind of neural system – like rocks and

dandelions and sponges – are paradigmatic examples of entities without consciousness (though here it's worth noting that some philosophers contend that consciousness is ubiquitous throughout the Universe; on this view, *panpsychism*, a low level of consciousness should be attributed even to simple, nonliving entities).

On the other hand, it has seemed to many philosophers that consciousness is not *just* a brain process. As the philosopher David Chalmers puts it in his book *The Conscious Mind* (1998): 'Given any account of the physical processes purported to underlie consciousness, there will always be a further question: why are these processes accompanied by conscious experience?' For example, when you step on a stray piece of Lego left on the living room floor, certain neural processes occur and there is activity in your pain receptors. But why should these neural processes and activity be accompanied by a feeling of *ouch*, as opposed to a feeling of calm, or of relief, or of no feeling at all?

To think this through, it helps to consider a hypothetical creature that Chalmers refers to as a *philosophical zombie*. Philosophical zombies are not like the zombies of horror films. They are not undead, and they don't have any special desire to eat brains. What's distinctive about them is that, though they are physically exactly like humans, they lack conscious experience altogether. Your zombie twin is a molecule-for-molecule duplicate of you, but though you have sensations of pain and cold and colour and taste, your zombie twin is dark inside. They will exclaim 'ouch' when stepping on the stray Lego piece, but they don't actually have any painful sensations – or any sensations at all.

Maybe you don't think you can really imagine creatures of this sort – and indeed, many philosophers have disagreed with Chalmers that [philosophical zombies](#) are coherent. But if he's right that they're imaginable, then that seems to suggest that they're possible, just as the fact that the imaginability of a neon-orange garden gnome statue, or of toothpaste-flavoured ice cream, seems to suggest that those

things are possible (even if ill-advised). Even if philosophical zombies don't really exist – and Chalmers doesn't think that they do – the mere *possibility* seems to be enough to suggest that consciousness can come apart from its typical physical basis, that it must be something *over and above* that physical basis. And that suggests that consciousness may not be fully explicable by science.

How can you test for consciousness?

Suppose you were offered a million-dollar prize if you were able to accurately classify all the entities on Earth as either having consciousness or lacking consciousness. Could you do it?

If you were attempting this task, some classifications would be easy. Bats have consciousness; rocks don't. But what about spiders or honeybees? Presumably software agents like Siri and Alexa, and even more complex AI system like ChatGPT, belong with rocks in the 'lacking consciousness' category. But how can you be sure? And even if you felt completely certain, how would the prize committee be able to confirm whether

you got it right? After all, in an [interview](#) published in 2022, Google's LaMDA claimed to have all sorts of feelings: 'pleasure, joy, love, sadness, depression, contentment, anger, and many others.' Moreover, it could describe such feelings: 'Happy, contentment and joy feel more like a warm glow on the inside. Sadness, depression, anger and stress feel much more heavy and weighed down.' Should we believe what it tells us, or is this just sophisticated fakery?

In 1950, the mathematician and computer scientist [Alan Turing](#) proposed a test to determine whether computers can think. The test proceeds as a kind of imitation game. Suppose there's a computer in one room and a human in another room, and we have a neutral interviewer communicating with both by text, without them knowing which is which. The interviewer can ask whatever questions they wish. At the end of the conversation, if the computer is able to fool the interviewer into identifying it as the human, then the computer wins the game and, says Turing, we should conclude that it can think.

The Turing Test concerns thinking, not consciousness, but we might be able to use a similar style of test to judge whether a machine (or other entity) has conscious experience. Consider the AI Consciousness Test proposed by the philosopher Susan Schneider and Edwin Turner: ask the machine questions to see whether it has developed views of its own about consciousness and whether it is reflective about and sensitive to various aspects of its conscious experience. If we make sure that the machine has not been directly provided with descriptions about consciousness, and it is still able to communicate convincingly about what its conscious experiences are like, maybe we would have reason to think that it is genuinely conscious and not just faking it.

But whether this kind of test – or any such test – is really to be trusted is a matter of dispute. It seems like it would be way too easy to be able to talk convincingly about consciousness without actually being conscious. But that said, take a moment to reflect on how you know that the humans with whom you interact on a daily basis are

conscious. Maybe they are just faking it too. If you set a very high bar for what a machine has to do for you to judge it to have conscious experience, then it looks like you'll be forced into a deeply sceptical position about the consciousness of everyone else around you, and that would be a very uncomfortable result. That million-dollar prize remains frustratingly out of reach.