

There Is No Such Thing as Conscious Thought

Philosopher Peter Carruthers insists that conscious thought, judgment and volition are illusions. They arise from processes of which we are forever unaware

• By [Steve Ayan](#)



Peter Carruthers, Distinguished University Professor of Philosophy at the University of Maryland, College Park, is an expert on the philosophy of mind who draws heavily on empirical psychology and cognitive neuroscience. He outlined many of his ideas on conscious thinking in his 2015 book *The Centered Mind: What the Science of Working Memory Shows Us about the Nature of Human Thought*. More recently, in 2017, he published a paper with the astonishing title of “The Illusion of Conscious Thought.” In the following excerpted conversation, Carruthers explains to editor Steve Ayan the reasons for his provocative proposal.

What makes you think conscious thought is an illusion?

I believe that the whole idea of conscious thought is an error. I came to this conclusion by following out the implications of the two of the main theories of consciousness. The first is what is called the Global Workspace Theory, which is associated with neuroscientists Stanislas Dehaene and Bernard Baars. Their theory states that to be considered conscious a mental state must be among the contents of working memory (the “user interface” of our minds) and thereby be available to other mental functions, such as decision-making and verbalization. Accordingly, conscious states are those that are “globally broadcast,” so to speak. The alternative view, proposed by Michael Graziano, David Rosenthal and others, holds that conscious mental states are simply those that you know of, that you are directly aware of in a way that doesn’t require you to interpret yourself. You do not have to read your own mind to know of them. Now, whichever view you adopt, it turns out that thoughts such as decisions and judgments should not be considered to be conscious. They are not accessible in working memory, nor are we directly aware of them. We merely have what I call “the illusion of immediacy”—the false impression that we know our thoughts directly.

One might easily agree that the *sources* of one’s thoughts are hidden from view—we just don’t know where our ideas come from. But once we have them and we know it, that’s where consciousness begins. Don’t we have conscious thoughts at least in this sense?

In ordinary life we are quite content to say things like “Oh, I just had a thought” or “I was thinking to myself.” By this we usually mean instances of inner speech or visual imagery, which are at the center of our stream of consciousness—the train of words and visual contents represented in our minds. I think that these trains are indeed conscious. In neurophilosophy, however, we refer to “thought” in a much more specific sense. In this view, thoughts include only nonsensory mental attitudes, such as judgments, decisions, intentions and goals. These are amodal, abstract events, meaning that they are not sensory experiences and are not tied to sensory experiences. Such thoughts never figure in working memory. They never become conscious. And we only ever know *of* them by interpreting what does become conscious, such as visual imagery and the words we hear ourselves say in our heads.

So consciousness always has a sensory basis? I claim that consciousness is always bound to a sensory modality, that there is inevitably some auditory, visual or tactile aspect to it. All kinds of mental imagery, such as inner speech or visual memory, can of course be conscious. We see things in our mind’s eye; we hear our

inner voice. What we are conscious of are the sensory-based contents present in working memory.

In your view, is consciousness different from awareness? That's a difficult question. Some philosophers believe that consciousness can be richer than what we can actually report. For example, our visual field seems to be full of detail—everything is just there, already consciously seen. Yet experiments in visual perception, especially the phenomenon of inattentional blindness, show that in fact we consciously register only a very limited slice of the world. [*Editors' note: A person experiencing inattentional blindness may not notice that a gorilla walked across a basketball court while the individual was focusing on the movement of the ball.*] So, what we think we see, our subjective impression, is different from what we are actually aware of. Probably our conscious mind grasps only the gist of much of what is out there in the world, a sort of statistical summary. Of course, for most people consciousness and awareness coincide most of the time. Still, I think, we are not directly aware of our thoughts. Just as we are not directly aware of the thoughts of other people. We interpret our own mental states in much the same way as we interpret the minds of others, except that we can use as data in our own case our own visual imagery and inner speech.

You call the process of how people learn their own thoughts interpretive sensory access, or ISA. Where does the interpretation come into play?

Let's take our conversation as an example—you are surely aware of what I am saying to you at this very moment. But the interpretative work and inferences on which you base your understanding are not accessible to you. All the highly automatic, quick inferences that form the basis of your understanding of my words remain hidden. You seem to just *hear* the meaning of what I say. What rises to the surface of your mind are the results of these mental processes. That is what I mean: The inferences themselves, the actual workings of our mind, remain unconscious. All that we are aware of are their products. And my access to your mind, when I listen to you speak, is not different in any fundamental way from my access to my own mind when I am aware of my own inner speech. The same sorts of interpretive processes still have to take place.

Why, then, do we have the impression of direct access to our mind?

The idea that minds are transparent to themselves (that everyone has direct awareness of their own thoughts) is built into the structure of our "mind reading" or "theory of mind" faculty, I suggest. The assumption is a useful heuristic when interpreting the statements of others. If someone says to me, "I want to help you," I have to interpret whether the person is sincere, whether he is speaking literally or ironically, and so on; that is hard enough. If I also had to interpret whether he is interpreting his own mental state correctly, then that would make my task impossible. It is far simpler to assume that he knows his own mind (as,

generally, he does). The illusion of immediacy has the advantage of enabling us to understand others with much greater speed and probably with little or no loss of reliability. If I had to figure out to what extent others are reliable interpreters of themselves, then that would make things much more complicated and slow. It would take a great deal more energy and interpretive work to understand the intentions and mental states of others. And then it is the same heuristic transparency-of-mind assumption that makes my own thoughts seem transparently available to me.

What is the empirical basis of your hypothesis?

There is a great deal of experimental evidence from normal subjects, especially of their readiness to falsely, but unknowingly, fabricate facts or memories to fill in for lost ones. Moreover, if introspection were fundamentally different from reading the minds of others, one would expect there to be disorders in which only one capacity was damaged but not the other. But that's not what we find. Autism spectrum disorders, for example, are not only associated with limited access to the thoughts of others but also with a restricted understanding of oneself. In patients with schizophrenia, the insight both into one's own mind and that of others is distorted. There seems to be only a single mind-reading mechanism on which we depend both internally and in our social relations.

What side effect does the illusion of immediacy have?

The price we pay is that we believe subjectively that we are possessed of far greater certainty about our attitudes than we actually have. We believe that if we are in mental state X, it is the same as being in that state. As soon as I *believe* I am hungry, I am. Once I *believe* I am happy, I am. But that is not really the case. It is a trick of the mind that makes us equate the act of thinking one has a thought with the thought itself.

What might be the alternative? What *should* we do about it, if only we could?

Well, in theory, we would have to distinguish between an experiential state itself on the one hand and our judgment or belief underlying this experience on the other hand. There are rare instances when we succeed in doing so: for example, when I feel nervous or irritated but suddenly realize that I am actually hungry and need to eat.

You mean that a more appropriate way of seeing it would be: "I think I'm angry, but maybe I'm not"?

That would be one way of saying it. It is astonishingly difficult to maintain this kind of distanced view of oneself. Even after many years of consciousness studies, I'm still not all that good at it (*laughs*).

Brain researchers put a lot of effort into figuring out the neural correlates of consciousness, the NCC. Will this endeavor ever be successful?

I think we already know a lot about how and where working memory is represented in the brain. Our philosophical concepts of what consciousness actually is are much more informed by empirical work than they were even a few decades ago. Whether we can ever close the gap between subjective experiences and neurophysiological processes that produce them is still a matter of dispute.

Would you agree that we are much more unconscious than we think we are?

I would rather say that consciousness is not what we generally think it is. It is not direct awareness of our inner world of thoughts and judgments but a highly inferential process that only gives us the impression of immediacy.

Where does that leave us with our concept of freedom and responsibility?

We can still have free will and be responsible for our actions. Conscious and unconscious are not separate spheres; they operate in tandem. We are not simply puppets manipulated by our unconscious thoughts, because obviously, conscious reflection does have effects on our behavior. It interacts with and is fueled by implicit processes. In the end, being free means acting in accordance with one's own reasons—whether these are conscious or not.

Briefly Explained: Consciousness

Consciousness is generally understood to mean that an individual not only has an idea, recollection or perception but also *knows* that he or she has it. For perception, this knowledge encompasses both the experience of the outer world (“it’s raining”) and one’s internal state (“I’m angry”). Experts do not know how human consciousness arises. Nevertheless, they generally agree on how to define various aspects of it. Thus, they distinguish “phenomenal consciousness” (the distinctive feel when we perceive, for example, that an object is red) and “access consciousness” (when we can report on a mental state and use it in decision-making).

Important characteristics of consciousness include subjectivity (the sense that the mental event belongs to me), continuity (it appears unbroken) and intentionality (it is directed at an object). According to a popular scheme of consciousness known as Global Workspace Theory, a mental state or event is conscious if a person can bring it to mind to carry out such functions as decision-making or remembering, although how such accessing occurs is not precisely understood. Investigators assume that consciousness is not the product of a single region of the brain but of larger neural networks. Some theoreticians go so far as to posit

that it is not even the product of an individual brain. For example, philosopher Alva Noë of the University of California, Berkeley, holds that consciousness is not the work of a single organ but is more like a dance: a pattern of meaning that emerges *between* brains. –S.A.

This article originally appeared in *Gehirn&Geist* and was reproduced with permission.

[Rights & Permissions](#)

MORE TO EXPLORE

The Opacity of Mind: An Integrative Theory of Self-Knowledge. Peter Carruthers. Oxford University Press, 2011.

The Centered Mind: What the Science of Working Memory Shows Us about the Nature of Human Thought. Peter Carruthers. Oxford University Press, 2015.

The Illusion of Conscious Thought. Peter Carruthers in *Journal of Consciousness Studies*, Vol. 24, Nos. 9–10, pages 228–252; 2017.