MIT Press • Open Encyclopedia of Cognitive Science

Consciousness

Claire Sergent[1] Frédérique de Vignemont[2]

[1]Integrative Neuroscience and Cognition Center, Université Paris Cité, CNRS,
[2]Institut Jean Nicod, ENS-PSL University, EHESS, CNRS

MIT Press

Published on: Jun 02, 2026

URL: https://oecs.mit.edu/pub/3daueu81

Consciousness refers to our subjective experience, broadly construed. It thus embraces a wide range of phenomena from wakefulness and sentience to what it is like to have a particular experience. Seeing the blueness of the ocean, hearing the sounds of the seagulls, and feeling pure joy: the subjective feelings that characterize these experiences seem hard to capture from a scientific viewpoint. One may then wonder in what sense these mental experiences differ at the cognitive and neural levels from those that are not conscious. This question is all the more important because, even when awake, an important part of the information processed by the brain escapes consciousness while still influencing cognition, decision-making, and behavior. Research on consciousness has flourished during recent decades, tackling some of the fundamental challenges associated with studying this inherently subjective material and bringing many important results as well as lively debates.

History

In the late 19th and early 20th century, psychology was conceived as a science of consciousness. Psychologists, and notably William James (1890), relied on attentive observation of his own mental activity to uncover the structure and properties of the stream of consciousness [see Introspection]. However, this introspective method has its limits. In particular, it can easily lead to contradictory conclusions among different experts, and it cannot guarantee objective results. Psychology then turned to behaviorism, which focused exclusively on observable behaviors and their relationship to external stimuli while expelling consciousness and inner mental states from the field of investigation (Watson, 1913).

With the rise of computationalism and cognitivism in the 1950s, it once again became scientifically legitimate to go beyond behavior, but it was primarily computations that were at the core of the cognitive interest, with no interest in their relation to subjective experiences. The issue of consciousness was left to philosophers. On the one hand, Thomas Nagel (1974) questioned the possibility to objectify what is essentially subjective [see Qualia]. In a famous article asking what it is like to be a bat, he argued that one would never be able to answer this question and concluded that consciousness makes the mind–body problem, that is, the relation between the brain and mentality, “hopeless.” On the other hand, Daniel Dennett (1992) objected that the difficulty for humans to imagine what it is like to be a bat does not prove the existence of a metaphysical property. He instead contended that a truly thorough description of a mental phenomenon in terms of all its behavioral, cognitive, and functional properties would not leave any phenomenon left to be explained. He provided influential arguments for bringing the theme of consciousness back into the field of scientific investigation by overcoming the apparent contradiction of building objective knowledge on what is an inherently subjective material. The study of consciousness was no longer conceived as hopeless.

It was the neuroscientist Bernard Baars (1989) who offered the first cognitive model of consciousness. In this model, conscious processing corresponds to the establishment of global communication across the

brain, allowing cooperation among specialized areas that would otherwise operate in an automatic mode. Since then, it has been well accepted that there is a need to integrate consciousness into the field of cognitive sciences and to refine the conceptual analysis of consciousness and the difficulties it raises. In 1995, two foundational papers appeared. Ned Block (1995) proposed a distinction between two types of consciousness, corresponding to the subjective character of conscious experiences, on the one hand, and to their cognitive accessibility and reportability on the other. In parallel, David Chalmers (1995) argued that the *hard problem* of consciousness is its subjective dimension, while the other problems, such as discriminative capacities, the mechanisms of attention, and the integration of information, were easier because they were only a matter of information processing.

Core concepts

The word "consciousness" can be used to mean many different things. Several distinctions are important for understanding contemporary debates.

Creature consciousness

First, one can ask whether a given system is capable of being conscious (or sentient) at all. This question is at the core of debates on infant consciousness (e.g., at which stage in development is an infant conscious?), animal consciousness (e.g., are insects and fish conscious?), and, more recently, on artificial intelligence (e.g., can a machine ever be conscious?) [see Infant Consciousness].

States of consciousness

Second, something that possesses creature consciousness can be in different states of consciousness. For humans, there may be different states of consciousness, such as wakefulness, sleep, coma, anesthesia, meditation, and so forth. At the conceptual level, one may ask whether these different states correspond to different *degrees* or *levels* of consciousness. At the experimental level, one can contrast brain activity across different states of consciousness (e.g., wakefulness vs. sleep).

Conscious mental activities

Third, within conscious systems that are in a certain state of consciousness one can ask whether a specific mental event—such as the perception of a word that appears only briefly on a screen—is conscious rather than not [see Unconscious Inference]. The challenge is to determine in which conditions a given content is available to consciousness and what types of content can be conscious. At the experimental level, one can then contrast the brain response to stimuli when participants report experiencing them consciously and when they do not, hopefully shedding light on the mechanisms that distinguish the two cases. A more recent line of investigation probes spontaneous conscious content by occasionally asking participants "what they have on their mind" and trying to relate brain activity with the different types of mental activity that they report (Smallwood & Schooler, 2015).

Reportability and heterophenomenology

The most straightforward benchmark to distinguish a conscious from a nonconscious state is the ability to report the content. This results in a very clean contrast between conscious and nonconscious processing in critical situations in which the stimulus is always the same, and only the subjective experience varies. Hence, reportability has been and still is an important operational criterion, even though it is not applicable for infant and animal consciousness.

Current approaches to reporting aim at avoiding the pitfalls associated with naïve approaches to introspection. According to what has been called the *method of heterophenomenology* (Dennett, 1992), subjective reports are considered from a third-person perspective. All investigators have access to the same data, and these data can be analyzed with objective methods.

Phenomenal overflow and phenomenal versus access consciousness

For some, reportability is a defining property of conscious processing (Cohen & Dennett, 2011), but others argue in favor of what has been called *phenomenal overflow*, according to which conscious experience can be richer than what one can report (Block, 2007). The hypothesis of *phenomenal overflow* finds its origin in studies using *partial recall* (Sperling, 1960). When participants are presented with three lines of four letters and asked to freely report the letters after the array has disappeared, they are able to report, on average, four letters. However, surprisingly it has been shown that they are potentially able to report *any* four letters of the array if, after the array has disappeared, they are presented a brief sound with different pitches corresponding to the different rows and asked to report only the letters of the corresponding row. This result has been taken as evidence that *iconic memory* has a larger capacity than *working memory*. It is also one of the main arguments in favor of the dissociation between two types of consciousness (Block, 1995). *Phenomenal consciousness* refers to the experiential properties to the inner subjective feel that seems to evade any functionalist or computationalist account. By contrast, *access consciousness* (also known as *cognitive reportability*; Block, 2007) refers to the capacity for a state to be used in reasoning, language, and the voluntary control of action. However, the issue of *phenomenal overflow* is still intensely debated. In particular, one may argue that it is conscious perception itself, and not only working memory, that can still be influenced by events occurring after the stimulus has disappeared (Sergent, 2018; Sergent et al., 2013).

Questions, controversies, and new developments

Identifying the neural correlates of consciousness

The neural correlates of consciousness have been defined as “the minimal set of neuronal events that gives rise to a specific aspect of a conscious percept” (Crick & Koch, 2003). A key method consists in contrasting brain activity associated with different subjective experiences (often inferred through report) with minimal or no change to the external stimulation. These experiments are carried out using protocols such as *visual* or *auditory masking* (a visual or auditory stimulus is presented very briefly and preceded

and followed by masking pictures or sounds), *inattentional blindness* (clearly visible stimuli can be missed if attention is directed towards other aspects of the scene), *binocular rivalry* (when two different pictures are presented to each eye, conscious perception spontaneously alternates between the two), and *attentional blink* (paying attention to a specific stimulus opens a critical time window of a few hundred of milliseconds during which conscious perception of subsequent stimuli is impaired). Across a wide range of different protocols and sensory modalities, it has been found that conscious perception is usually accompanied with stronger activation in the sensory cortices associated with the specific stimulus, with the implication of a wider network of areas, usually beyond sensory cortices, and with changes in communication across brain areas (Sergent & Naccache, 2012).

However, the interpretation of these results is still a matter of debate. Notably, it is unclear what to conclude about the precise mechanisms of consciousness on the basis of these neural correlates. Indeed, the contrast between conscious and nonconscious does not result exclusively from the processes associated with consciousness. It can also include additional processes that correlate with them. For instance, spontaneous fluctuations before the stimulus appears or in the early stages of stimulus processing might favor whether it becomes conscious or not (Aru et al., 2012; Sergent & Naccache, 2012). Furthermore, because participants are typically asked to report their perception of the stimulus, the neural correlates obtained might confound conscious processing per se with postperceptual activities linked with answering the question. Current developments therefore aim at addressing these confounds, notably through *no-report* paradigms in which participants do not perform a task on the stimulus of interest (Tsuchiya et al., 2015).

Mechanisms underlying consciousness

Several cognitive and neural models of consciousness have been proposed. According to g*lobal workspace theory*, a content is conscious if it is globally available to other cognitive processes, notably attention, cognitive control, decision-making, and action (Baars, 1989; Dehaene & Changeux, 2011). This model has been mostly dedicated to understanding conscious processing in the context of decision-making. Hence, although it can explain cognitive accessibility, it is not clear how it can account for the subjective character of consciousness. A recent addition to the model, the *global playground*, extends this proposition to describe conscious processing even in the absence of a cognitive task. What makes a content conscious is its global availability to other cognitive processes, but these processes need not be restricted to cognitive control or decision-making; they include emotional and motivational processing and an incorporation of this content into episodic memory and into processes associated with the sense of self. This extended notion of cognitive availability relates more directly to the subjective aspects of conscious processing (Sergent et al., 2021).

Another major trend situates the origin of consciousness in higher-order mechanisms of self-monitoring (Lau & Rosenthal, 2011) [see Self-Consciousness]. According to these *higher-order theories*, a first-order content is conscious only if it is the object of a second-order representation; for example, a visual stimulus

such as a tilted line is conscious only if first-order information in the visual cortex (tilted line) is re-represented in higher-level regions in a metacognitive form akin to, "I see a tilted line." It may seem, however, that on this view, consciousness then requires relatively sophisticated cognitive capacities, which may not be available to young infants and to species with simpler neural systems.

In contrast, *the local recurrency hypothesis* is less cognitively demanding. It postulates that conscious perception is associated with the establishment of local recurrent loops among sensory areas during the early phases of sensory processing, independently of attentional or higher-order processes (Lamme, 2006).

Finally, according to *integrated information theory*, conscious processing is associated with a specific type of interaction among entities, such as neurons within neural populations, whereby the state of each entity is influenced by all the other entities so that the system is highly integrated (Tononi, 2012). In this view, any system with connected entities can be attributed a measure (*phi*) of their integration, which is then taken to be a measure of consciousness.

The function of consciousness

One of the main mysteries about *phenomenal consciousness* is whether it has any function and, if so, what. This question is illustrated by the following thought experiment: Can one conceive that there are zombies that are strictly identical to humans at the behavioral, neural, and even chemical level but that lack conscious experiences (Chalmers, 1996)? If the answer is positive, then consciousness is *epiphenomenal*; it plays no functional role. However, if one can show that the notion of zombie is not conceptually coherent, then it means that consciousness makes a difference [see Qualia].

The evolutionary origins of consciousness

From an evolutionary perspective, there are two major questions. First, there is a debate over whether consciousness evolved gradually or appeared suddenly. Second, there is disagreement about whether consciousness evolved along a single path or multiple parallel paths, implying that it might take different forms depending on different organisms (Godfrey-Smith, 2017).

Is consciousness an illusion?

In the past decade, some philosophers have suggested that the mystery is not so much consciousness itself but rather the fact that we believe in consciousness. According to the position known as *illusionism*, one is mistaken in believing that there is something it is like to undergo experiences. On this view, mental states have no phenomenal properties (Frankish, 2016; Kammerer, 2021). More fundamentally, illusionism raises questions about the intuitions that are commonly shared about consciousness and about their origin.

Broader connections

Research on consciousness has many practical implications, notably in clinics; current knowledge on the neural signatures of consciousness is starting to be used to probe consciousness in patients who are noncommunicating (Engemann et al., 2018), to better understand the effects of psychedelics and anesthetics (Ferrarelli et al., 2010), and to investigate conscious processing in certain psychiatric conditions (Berkovitch et al., 2021). Consciousness is also associated with major ethical considerations. Indeed, conscious beings are typically held to be significantly more morally valuable than nonconscious ones, although the value of consciousness remains to be determined. Therefore, the question of consciousness is at the heart of debates about animal rights. More recently, the successes of artificial intelligence raise the question of whether artificial consciousness is possible [see Consciousness and AI]. Finally, the notion of consciousness connects directly with the notion of moral responsibility, central to the functioning of our societies.

Acknowledgments

We thank Jérôme Sackur for sharing with us his insightful views.

Further reading

- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*(2), 227-247. https://doi.org/10.1017/S0140525X00038188
- Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6*(2), 119-126. https://doi.org/10.1038/nn0203-119
- Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 200-227. https://doi.org/10.1016/j.neuron.2011.03.018
- Dennett, D. C. (1992). *Consciousness explained*. Penguin.

References

- Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience & Biobehavioral Reviews*, *36*(2), 737-746. https://doi.org/10.1016/j.neubiorev.2011.12.003 ↩
- Baars, B. J. (1989). *A cognitive theory of consciousness*. Cambridge University Press. ↩

- Berkovitch, L., Charles, L., Del Cul, A., Hamdani, N., Delavest, M., Sarrazin, S., Mangin, J.-F., Guevara, P., Ji, E., Albis, M.-A., Gaillard, R., Bellivier, F., Poupon, C., Leboyer, M., Tamouza, R., Dehaene, S., & Houenou, J. (2021). Disruption of conscious access in psychosis is associated with altered structural brain connectivity. *Journal of Neuroscience*, *41*(3), 513-523. https://doi.org/10.1523/JNEUROSCI.0945-20.2020

  ↩

- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*(2), 227-247. https://doi.org/10.1017/S0140525X00038188

  ↩

- Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and brain sciences*, *30*(5-6), 481-499. https://doi.org/10.1017/S0140525X07002786

  ↩

- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*(3), 200–219.

  ↩

- Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory* (2nd ed.). Oxford University Press.

  ↩

- Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, *15*(8), 358-364. https://doi.org/10.1016/j.tics.2011.06.008

  ↩

- Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6*(2), 119-126. https://doi.org/10.1038/nn0203-119

  ↩

- Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 200-227. https://doi.org/10.1016/j.neuron.2011.03.018

  ↩

- Dennett, D. C. (1992). *Consciousness explained*. Penguin.

  ↩

- Engemann, D. A., Raimondo, F., King, J.-R., Rohaut, B., Louppe, G., Faugeras, F., Annen, J., Cassol, H., Gosseries, O., Fernandez-Slezak, D., Laureys, S., Naccache, L., Dehaene, S., & Sitt, J. D. (2018). Robust EEG-based cross-site and cross-protocol classification of states of consciousness. *Brain*, *141*(11), 3179-3192. https://doi.org/10.1093/brain/awy251 ↩
- Ferrarelli, F., Massimini, M., Sarasso, S., Casali, A., Riedner, B. A., Angelini, G., Tononi, G., & Pearce, R. A. (2010). Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. *Proceedings of the National Academy of Sciences*, *107*(6), 2681-2686. https://doi.org/10.1073/pnas.0913008107 ↩
- Frankish, K. (2016). Illusionism as a theory of consciousness. *Journal of Consciousness Studies, 23*(11-12), 11–39. ↩
- Godfrey-Smith, P. (2017). The evolution of consciousness in phylogenetic context. In K. Andrews & J. Beck (Eds.), *The Routledge handbook of philosophy of animal minds* (pp. 216–226). Routledge. ↩
- James, W. (1890). *The principles of psychology*. Henry Holt and Company. ↩
- Kammerer, F. (2021). The illusion of conscious experience. *Synthese*, *198*(1), 845-866. https://doi.org/10.1007/s11229-018-02071-y ↩
- Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*(11), 494-501. https://doi.org/10.1016/j.tics.2006.09.001 ↩
- Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, *15*(8), 365-373. https://doi.org/10.1016/j.tics.2011.05.009 ↩
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435–450. https://doi.org/10.2307/2183914 ↩
- Sergent, C. (2018). The offline stream of conscious representations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1755), 20170349. https://doi.org/10.1098/rstb.2017.0349 ↩

- Sergent, C., & Naccache, L. (2012). Imaging neural signatures of consciousness: "what," "when," "where" and "how" does it work? *Archives Italiennes de Biologie*, *150*(2-3), 91-106. https://doi.org/10.4449/aib.v150i2.1270

  ↩

- Sergent, C., Corazzol, M., Labouret, G., Stockart, F., Wexler, M., King, J.-R., Meyniel, F., & Pressnitzer, D. (2021). Bifurcation in brain dynamics reveals a signature of conscious processing independent of report. *Nature Communications*, *12*(1), 1149. https://doi.org/10.1038/s41467-021-21393-z

  ↩

- Sergent, C., Wyart, V., Babo-Rebelo, M., Cohen, L., Naccache, L., & Tallon-Baudry, C. (2013). Cueing attention after the stimulus is gone can retrospectively trigger conscious perception. *Current Biology*, *23*(2), 150-155. https://doi.org/10.1016/j.cub.2012.11.047

  ↩

- Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, *66*, 487-518. https://doi.org/10.1146/annurev-psych-010814-015331

  ↩

- Sperling, G. (1960). The information available in brief visual presentation. *Psychological Monographs: General and Applied, 74*(11), 1–29. https://doi.org/10.1037/h0093759

  ↩

- Tononi, G. (2012). Integrated information theory of consciousness: An updated account. *Archives Italiennes de Biologie*, *150*(2-3), 56-90. https://doi.org/10.4449/aib.v149i5.1388

  ↩

- Tsuchiya, N., Wilke, M., Frassle, S., & Lamme, V. A. (2015). No-report paradigms: Extracting the true neural correlates of consciousness. *Trends in Cognitive Science*, *19*(12), 757-770. https://doi.org/10.1016/j.tics.2015.10.002

  ↩

- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological review*, *20*(2), 158-177. https://doi.org/10.1037/h0074428

  ↩