

Understanding understanding: Is consciousness impossible to comprehend?

180201222
April 15, 2024

“Consciousness is a fascinating but elusive phenomenon: it is impossible to specify what it is, what it does, or why it has evolved. Nothing worth reading has been written on it.” - Stuart Sutherland, *The Macmillan Dictionary of Psychology* (1989)

Consciousness may be crudely defined as *what it is like* to be like something. I know that I am conscious, as I know what being me is like (Descartes’ “*cogito ergo sum*”). But how would I know that you, the reader, are conscious? Through conscious thought I can ask questions. Where does consciousness come from? When did consciousness evolve? Why are we even conscious in the first place? Do these questions even have answers? The rather pessimistic viewpoint presented above suggests not, however, these questions have persisted in the minds of philosophers and, more recently, scientists - who take a modern approach to answering them. This article will evaluate historical and modern arguments of the conscious mind, and assess the consensus among current researchers.

There is no universally agreed-upon definition of consciousness. The word originated from the Latin *con* (“together”) and *scio* (“to know”) - implying a shared knowledge, such as a secret [1]. The modern definition of consciousness may be attributed to John Locke, who describes consciousness as “the perception of what passes in a Man’s own mind” [2]. Thus, it is a form of awareness, not of the outside world as granted to us by our senses, but of one’s inner self.

The origin of conscious thought is not known. Cartesian duality is the idea that a conscious being consists of a soul which drives the body. The so-called “ghost in the shell” argument combines the physical and metaphysical to explain how crude collections of matter can experience thought, feeling and emotion [3]. But is the mind metaphysical, or could it be a consequence of specific patterns of behaviour in inert, thoughtless matter? Where does the intangible soul come from? There is an unquestionable connection between the mind and the brain, and neurological research like that of Francis Crick and Christof Koch in the 1990s aims to reveal the nature of this relationship. The motivation of this research lies in better understanding clinical medical cases, animal consciousness, and perhaps even machine consciousness.

Reportability is a vital tool for probing consciousness. Assuming that other consciousnesses exist outside of our own minds,¹ we are most confident if they are able to communicate their thoughts and feelings verbally or through some form of behaviour [5]. However, these *behavioural correlates* are not totally dependable. A sleepwalker shows

behaviours of a conscious person, yet may be completely void of conscious thought. Conversely, someone lying motionless could be experiencing vivid dreams.

Consciousness can be thought of as a spectrum, with levels which are experienced in different mental states. *Connected* consciousness describes the experiences associated with being awake, making decisions and perceiving one’s surroundings. Conversely, *disconnected* consciousness is experienced when asleep, but not unconscious - such as during a dream [6]. The transitions between these states provide excellent test beds for the brain patterns associated with consciousness, as they highlight the moments of both loss and recovery of consciousness [7]. By measuring electrical signals in the brain, neurologists can relate specific signal patterns to feelings such as hunger. These *neural correlates* are used by scientists to link activity in the brain with behaviour. Conscious thought has been constrained to an area of the brain known as the posterior cortex using neural correlates [8], however, it is still not known *why* certain brain patterns correlate to specific behaviours.

How can one perceive their favourite red roses, or a song which makes them cry, through nothing but a series of electrical signals travelling through brain matter comprised of unliving particles? How do we bridge the *explanatory gap* between the experiences and behaviours which we know of and can explain scientifically, with the unobservable nature of the conscious mind? This is known as the “hard problem” of consciousness - made famous by David Chalmers [9, 10].

Panpsychism is a theory which has been around since antiquity. Advocates of the theory, such as Philip Goff, believe that there is inherent consciousness in all matter, such that even a single proton would have the simplest form of awareness. Plants have a more complex form of experience, fish more complex still, mice, apes, and so on. In more complicated beings like humans, more complex consciousness arises [11]. The failure of panpsychism to explain how simple consciousnesses combine to form more complex experiences has made it an unpopular theory in modern philosophy, but it is not disregarded completely [12]. A variation of this theory called Integrated Information Theory came from Giulio Tononi. Here, consciousness is present where information is swapped via some mathematically defined mechanism. Not everything is conscious in this model, however the levels of consciousness are ambiguous. It was pointed

¹the Solipsism problem. The *Boltzmann brain paradox* is an interesting thought experiment on this [4].

out in a conference that a simple compact disc running error calculations would be far more conscious than a human under this definition [13].

The Global Neural Workspace theory argues that consciousness is achieved when information is shared between multiple specialised systems. Information is posted onto a “blackboard” of limited space, where all areas of the brain can access it. For example, a student consciously taking lecture notes would be unconsciously scribbling onto paper, without needing to think about forming each letter. If only local circuits have access to information (there is not enough room on the blackboard), the task is performed unbeknownst to the rest of the system [8]. In this framework, computer consciousness would become possible when artificial intelligence evolves more specialised circuits to deal with more tasks, which can independently access shared information.

Could an artificial intelligence become fully conscious before we have a definition for it? Goff mentions that is is hard to make progress on understanding consciousness because it defies observation [14]. Physicists can study the unobservable, such as fundamental particles or quantum phenomena, through theory. Vitally however, they do so to explain what *is* observed. As consciousness cannot be publicly observed, Goff argues, it is completely opaque to the scientific gaze. Nonetheless, scientific research into consciousness has accelerated in the last two decades, particularly in the field of neuroscience, following the efforts of Critch and Koch in the 1990s. Philosophers also continue to deliver new and evolved theories of the mind, providing frameworks in which scientists can work. The breakthrough was thought to be close once before, however.

In 2023, a 25-year-old bet between neuroscientist Koch and philosopher Chalmers was lost by the former, who believed there would be a clear consensus of consciousness by this time [13]. There are some, such as Goff, who believe we are no closer to a unified theory, but others believe we are not too far away. Upon conceding the bet, Koch doubled-down, betting that by 2048 we would understand the conscious mind. His unperturbed confidence was met by equal scepticism, with Chalmers accepting the bet, stating “I hope I’m wrong, but I suspect I’m right”. Chalmers is not of the same mind as Sutherland however, having stated that he believes we will have a solution to his famous hard problem within 100 years [10].

The conscious mind is the subject of extreme division between many philosophers and scientists. However, there is an opinion which unites all parties - that humans *will* one day understand it. Combining the efforts of both factions is key to unifying theories of the mind, and bridging the explanatory gap. What would be the applications of under-

standing the nature of consciousness? How would we perceive animals, plants, computers, and the world around us? Would we treat others differently? Would it simply become another fact of life, which is taught in primary schools with little-to-no fanfare? For now the mind can only wonder...

References

- [1] C. S. Lewis, *Studies in Words*. Cambridge University Press, 1990.
- [2] J. Locke, *An Essay Concerning Human Understanding*. Kay & Troutman, 1847.
- [3] W. G. Lycan, *Consciousness*. A Bradford Book, Cambridge, Mass. ; London: MIT Press, 1987.
- [4] TED-Ed, “The Boltzmann brain paradox - Fabio Pacucci,” 2022.
- [5] G. Tononi and C. Koch, “Consciousness: Here, there and everywhere?,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, p. 20140167, May 2015.
- [6] C. G. Brown, Emery, “Consciousness Is a Continuum, and Scientists Are Starting to Measure It,” 2024.
- [7] C. Koch, “Safely Switching Consciousness Off and On Again,” 2012.
- [8] C. Koch, “What Is Consciousness?,” 2018.
- [9] E. Thompson, *The Problem of Consciousness: New Essays in Phenomenological Philosophy of Mind*. No. 29 in Canadian Journal of Philosophy. Supplementary Volume, Calgary: University of Calgary Press, 2003.
- [10] Serious Science, “Hard Problem of Consciousness — David Chalmers,” 2016.
- [11] G. Cook, “Does Consciousness Pervade the Universe?,” 2020.
- [12] D. Falk, “Is Consciousness Part of the Fabric of the Universe?,” 2023.
- [13] J. Horgan, “A 25-Year-Old Bet about Consciousness Has Finally Been Settled,” 2023.
- [14] P. Goff, “Understanding Consciousness Goes Beyond Exploring Brain Chemistry,” 2023.