NOTEWORTHY, PHILOSOPHY

THE GREAT CONSCIOUSNESS DEBATE

13.03.2025

The most recent discussion abut consciousness was going in circles, having now a dozen of competing theories as summarized by chatGPT and reworked by me for omissions. Not sure if all the summaries and references are good, but maybe this is a good starting point to understand the background of a new paper.

1. 1. Global Workspace Theory (GWT) proposed by Bernard Baars.
   – Suggests that consciousness arises when information is “broadcast” to a global workspace in the brain, allowing different cognitive processes to access and use it.
   – Compares the brain to a theater, where only information in the spotlight of attention becomes conscious.
   – Baars, B. J. (1988). A Cognitive Theory of Consciousness. Cambridge University Press.
2. Integrated Information Theory (IIT) developed by Giulio Tononi.
   – Proposes that consciousness is linked to the amount of integrated information (Φ) in a system.
   – Suggests that the more a system is capable of integrating information across different parts, the more conscious it is.
   – Tononi, G. (2004). An information integration theory of consciousness. BMC Neuroscience, 5(1), 42.
3. Higher-Order Theories (HOT)**
   – Suggests that consciousness arises from thoughts about thoughts.
   – A mental state becomes conscious when there is a higher-order representation of that state in the brain.
   – Championed by researchers like David Rosenthal and Hakwan Lau.
   – Rosenthal, D. M. (2005). Consciousness and Mind. Oxford University Press.
4. Attention Schema Theory (AST) proposed by Michael Graziano.
   – Suggests that the brain creates a simplified model (a schema) of its own attention processes, which gives rise to subjective awareness.
   – Consciousness is a type of self-monitoring mechanism.
   – Graziano, M. S. A. (2013). Consciousness and the Social Brain. Oxford

University Press.

5. Predictive Processing & Bayesian Brain Theories
   – Suggests that the brain is constantly generating predictions about sensory input and updating them based on new information.
   – Consciousness arises from how well the brain models the world and resolves prediction errors.
   – Friston, K. (2010). The free-energy principle: a unified brain theory? Nature Reviews Neuroscience, 11(2), 127-138.
6. Quantum Theories of Consciousness like Penrose & Hameroff's Orchestrated Objective Reduction (Orch-OR) – suggest that consciousness arises from quantum processes in microtubules in brain neurons.
   – Highly controversial, as many neuroscientists doubt quantum effects play a significant role in brain function.
   – Penrose, R. (1989). The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics. Oxford University Press.
7. Panpsychism
   – Suggests that consciousness is a fundamental feature of the universe, present in all matter at some level.
   – Variants include **Integrated Information Panpsychism** (which aligns with IIT) and **Cosmopsychism** (which suggests the universe itself is conscious).
   – Philosophers like Philip Goff and Galen Strawson support versions of this idea.
   – Strawson, G. (2006). Realistic monism: Why physicalism entails panpsychism. Journal of Consciousness Studies, 13(10-11), 3-31.
8. Enactive and Embodied Theories
   – Consciousness is seen as emerging from the interaction between the brain, body, and environment.
   – Rather than being purely brain-based, it depends on active engagement with the world.
   – Varela, Thompson, and Noë have been key contributors.
   – Varela, F. J., Thompson, E., & Rosch, E. (1991). The Embodied Mind: Cognitive Science and Human Experience. MIT Press.
9. Recurrent Processing Theory (RPT) proposed by Victor Lamme.
   – Suggests that consciousness arises from recurrent (looping) activity in cortical networks.
   – When neural activity is only feedforward, it remains unconscious.
   – Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. Trends in Cognitive Sciences, 10(11), 494-501.
10. Interactionist Dualism Theory (IDT) proposed by Eccles and Popper
    – consciousness has a non-physical reality that interacts with the brain.

– influences the brain by intervening at specific synapses
– unlike materialist views, this theory suggests that the mind and brain are distinct but interact causally.
– criticized for lacking a clear mechanism for how a non-physical mind could affect a physical brain.
– Popper, K. R., & Eccles, J. C. (1977). The Self and Its Brain. Springer.

11. Recurrent Processing Theory (RPT) proposed by Victor Lamme.
– Suggests that consciousness arises from recurrent (looping) activity in cortical networks.
– When neural activity is only feedforward, it remains unconscious.
12. Orchestrated Objective Reduction (Orch-OR) proposed by Penrose and Hameroff
– suggests that consciousness arises from quantum processes within microtubules in brain neurons.
– claims that quantum coherence and reduction of wavefunctions within microtubules play a fundamental role in conscious experience.
– highly controversial
– Hameroff, S., & Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. Mathematics and Computers in Simulation, 40(3-4), 453-480.

No single theory is universally accepted. Some researchers believe that an integrated approach combining multiple theories may be necessary to fully explain consciousness. I am still intrigued by John John Eccles and Karl Popper book that I read in 1989 which may still be valid although containing “unscientific” metaphysical elements.

A new paper now just identifies the unscientific elements in he IIT / integrated information theory

One well-known proposal — integrated information theory — has recently been labeled as ‘pseudoscience’, which has caused a heated open debate. Here we discuss the case and argue that the theory is indeed unscientific because its core claims are untestable even in principle.

While I agree that IIT is not scientific for limited empirical support, I have doubts if this is a sufficient cause to deny the validity of the argument. Also metapyhsically *a priori* methods are justifed even if the rely only “on rational intuition and abstract reasoning from general principles rather than sensory experience.”

Independent of that, I am sharing the concerns of the authors if some untested theories like IIT is leading to serious consequences

> By promoting to the general public the untestable idea that 'Φ = consciousness', proponents of IIT may ultimately have an unjustified effect on law and policy, including on decisions that involve measures of quality of life, clinical triage, abortion, the rights of non-responsive patients, and welfare considerations for insects (?), organoids and artificial intelligences.