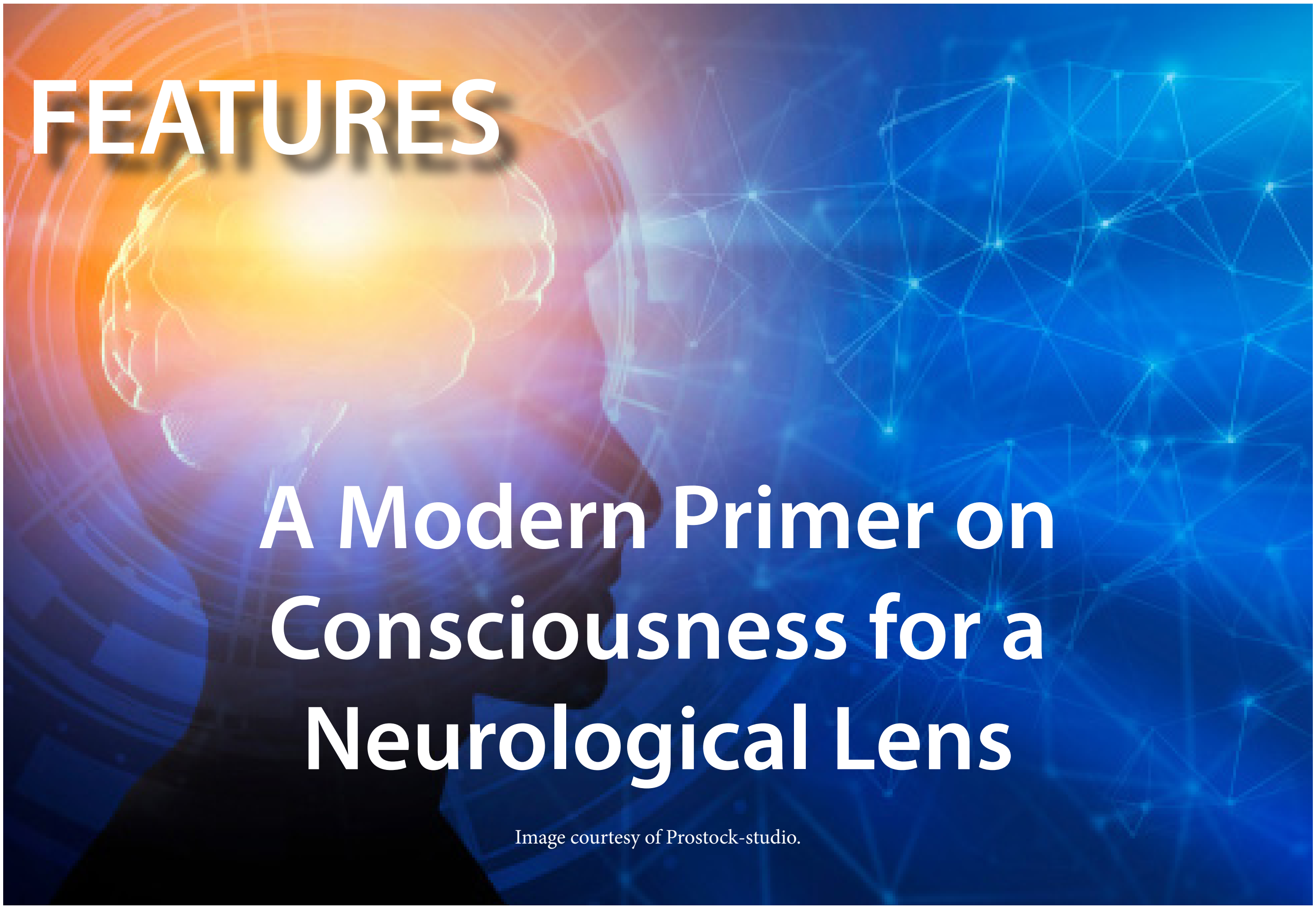

A Modern Primer on Consciousness for a Neurological Lens

Image courtesy of Prostock-studio.

Avery Mizrahi '28

Introduction

What does it mean to be human? Efforts to answer this question almost always end up mentioning one key term, popular and numerous in its definitions: consciousness. In day-to-day language, we might mention someone being knocked unconscious. Or we might say we were unconscious of our actions, such as driving home (in this case, interchangeable with unaware). For many philosophers of mind, the most succinct definition of consciousness is "what it is like" to be a human: unified, directed, and self-aware. It is the redness of an apple, or the painfulness of pain (Nagel, 1974).

Scientists attempt to answer the question of human meaning by formulating a biological basis of consciousness, which is paramount to understanding how human experience of life changes over time, occasionally for the worse. How does our consciousness change with age and neurodegenerative disease (NDD), when we lose the memories or functions that make us, us? How about after traumatic brain injury (TBI), when personality erodes in the face of neuronal loss? Of great relevance is the question of how we can clarify, quantify, and restore consciousness for patients afflicted by ailments such as these. In this article, I will briefly summarize how historical and modern theories of consciousness have been conceptualized and quantified with the goal of better treating neurological disorders.

Historical Foundations of Consciousness

As early as the 5th century BCE, the Buddha's foundational theory of consciousness described no consistent self, but rather the shifting impermanent instances that make up consciousness (Bodhi, 2000). Later, western philosophers tended to reject this view, and John Locke is credited with postulating consciousness as the consistent awareness of

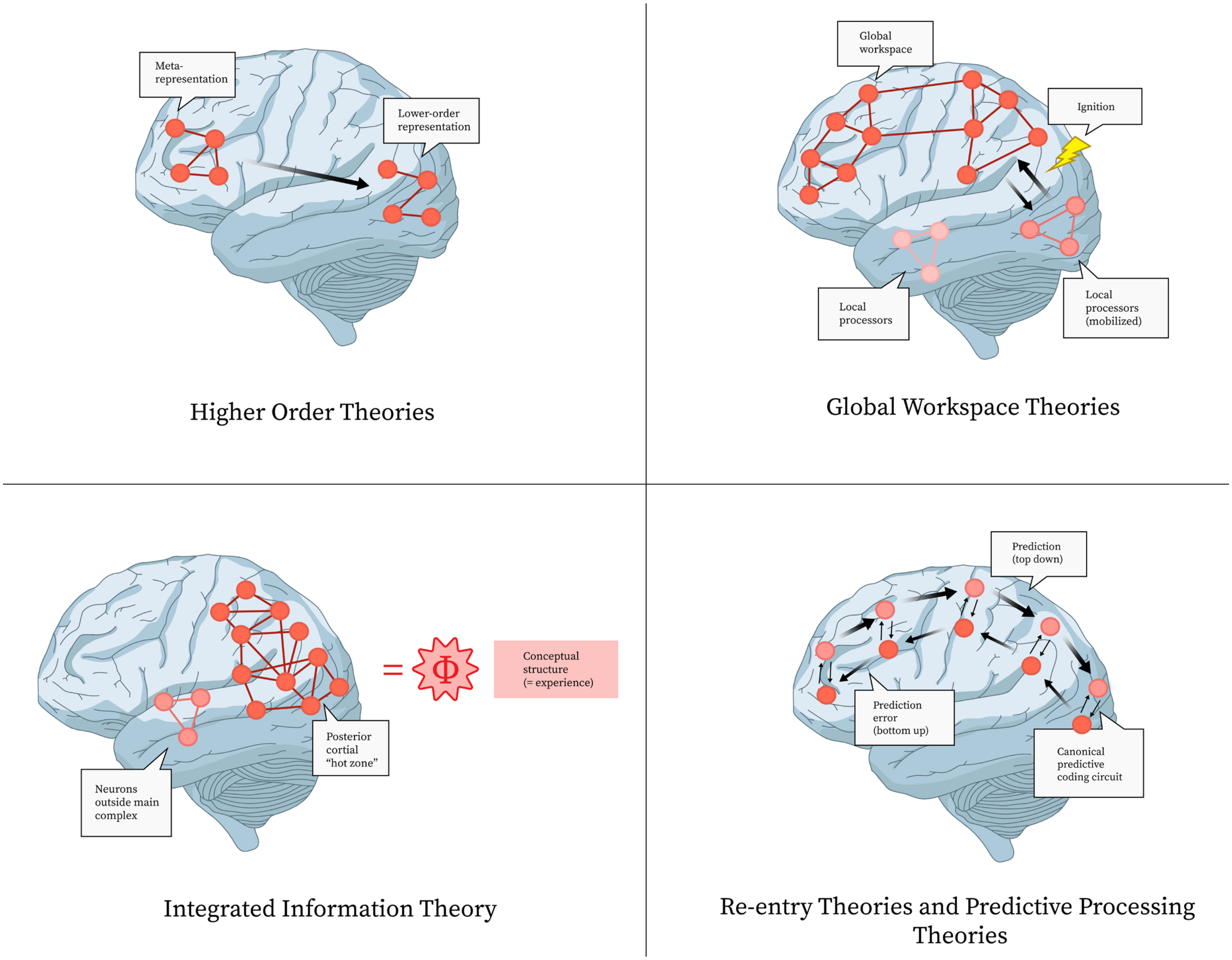

Figure 1. Anatomical networks involved in a few specific theories of consciousness (Adapted from Seth & Bayne, 2021).

thought: "I do say he can not think at any time, waking or sleeping, without being sensible of it" (Locke, 1689). However, as we now know—and as G.W. Leibniz posited as early as 1720—unconscious cognitive processes frequently bypass our conscious awareness; this is the example of being unaware of driving home, or sleeping. In 1787, Immanuel Kant modified this idea, arguing that consciousness is not the awareness or unawareness of thought, but rather the experience of a self situated in the world (Kant, 1787/1998). Kant inspired the Phenomenology movement of the early 20th century, which was characterized by studying the philosophy of lived experience (Smith, 2013). Modern theories elaborate on these philosophical notions in an attempt to create a biological hypothesis of how consciousness arises.

Modern Theories of Consciousness

The theories surrounding how the brain creates a subjective experience are numerous and contradictory. For example, the Attention Schema Theory (AST) argues the perception of self arises from the brain being inundated with information. Because of the sheer amount of information we take in, our attention is selective; therefore, consciousness is our awareness of selective attention (Graziano & Webb, 2015). Awareness of attention creates a sense of self, the "I" in the statement "I am having the conscious experience of this red apple." In the same way, our brain simplifies visual information and creates an internal representation of the outside world, AST argues our brain simplifies attention and creates an internal representation of ourselves. Therefore, AST may be beneficial in comparing how representations

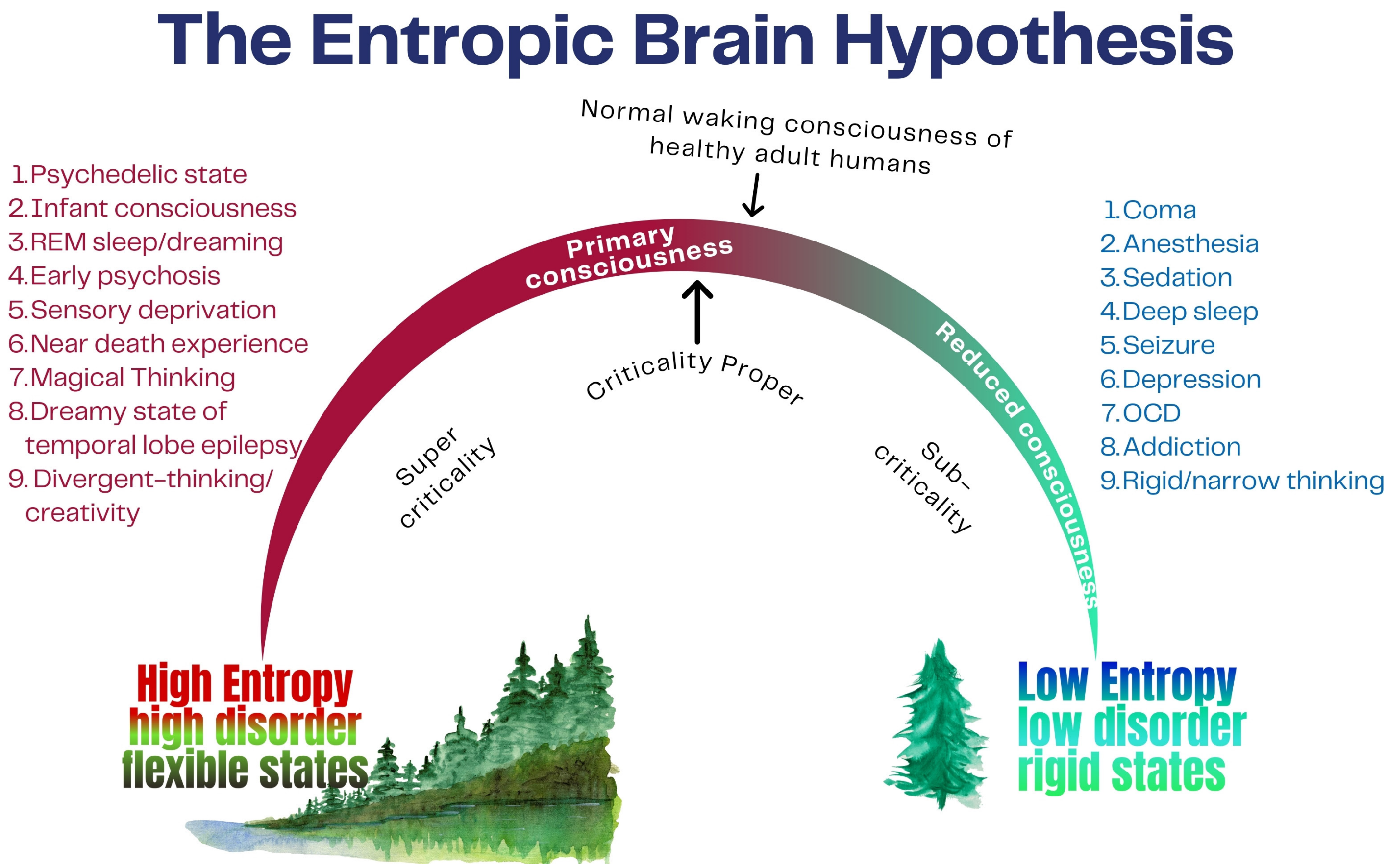

Figure 1. Various cognitive states and disorders may differ in entropic levels (Carhart-Harris, 2014).

of the self shift, from when memory deteriorates in Alzheimer's disease (AD) to when worldview changes in those suffering from Major Depressive Disorder.

A more mathematical approach to consciousness is the Integrated Information Theory (IIT), which proposes that one's degree of phenomenal experience can be translated into a single value (Seth & Bayne, 2021). This value is derived from the connectivity between each node (discrete brain regions) in a network where consciousness arises. IIT may also help identify the neural correlates of consciousness, or specific and minimal neural events that are associated with conscious experience (Dehaene et al., 2011). IIT is hotly debated because opponents state the brain is too complex to create one consistent value that adequately encompasses consciousness. The same researchers have created the more accessible Perturbational Complexity Index, which calculates consciousness values in a variety of states with reduced computational demand from electroencephalogram (EEG) data (Casali et al., 2013).

Another useful hypothesis that provides a cohesive framework for varieties of conscious experience is the Entropic Brain Hypothesis. This notion defines a conscious state by its entropy, or the flexibility of neural patterns (Carhart-Harris, 2014). While examples of flexible states on the "high entropy" side of the spectrum are psychedelic experiences or rapid eye movement (REM) sleep, examples of "low entropy" states are comas or Obsessive-Compulsive Disorder.

Other popular theories tend to differ on whether consciousness arises from higher-order networks (Higher Order Theories), a global consciousness hub (Global Workspace Theory), or the brain's attempt to minimize prediction errors (Predictive Processing Theory) (Seth & Bayne, 2021). Though the above theories may influence how we approach the underlying biological mechanisms of consciousness, some argue that theories alone are unhelpful without additional tools of measurement. Therefore, theorists work in conjunction with experiments and technology to measure and quantify consciousness, especially in diseased states.

Measurement and Quantification

A meaningful method for measuring consciousness must reliably detect both consciousness and its absence. One route taken is the Mirror Test for self-recognition (Dehaene, 2016). This intuitive test often consists of marking the forehead of an animal and seeing whether they wipe away the chalk when looking in the mirror. The theory proposes that if an animal can recognize itself in the mirror, there must be some sense of self-awareness. Interestingly, this behavior does not emerge in humans until 15-24 months of age, and is linked to introspection, mental states, and empathy. At the opposite end of life, experts have observed that self-recognition deteriorates with cognitive decline (Biringer & Anderson, 1992).

“A meaningful method for measuring consciousness must reliably detect both consciousness and its absence”

Figure 3. A puppy likely doesn’t recognize itself in a mirror (iStock Emma Jocelyn).

As appealing as the recognition test’s binary answer to consciousness might be, it is commonly thought that consciousness is a gradient rather than a question of all-or-none. To understand this spectrum, many neuroscientists have utilized tools such as functional Magnetic Resonance Imaging (fMRI) to look at activity and connectivity in the brain. Most commonly, the default mode networks (DMN), discovered from task-free resting state fMRI, have been identified as quantifiable features that vary in different states of consciousness (Raichle & Mintun, 2006). However, fMRI is not a perfect tool; it has high spatial resolution, but low temporal resolution (i.e., you can accurately view many regions of the brain, but not over a span of time). By comparison, electroencephalograms (EEG) have the opposite limitations, as a temporally, but not spatially, specific tool of measurement. Combined, these tools create an experimental paradigm to study consciousness both temporally and spatially.

Neurological Implications

Using the quantification and measurements of consciousness, scientists can observe how certain neural states change in the disordered brain, such as in NDDs. For example, early AD is correlated with decreased functional connectivity in the DMN during resting-state fMRI, suggesting an altered consciousness as the disease progresses (Simic et al., 2014).

Additional work surrounds how the brain recovers from coma to the normal function associated with wakefulness. Many potential neuronal activity states are associated with wakefulness, so neuroscientist Diany Calderon speculates that the brain may pass through different activity states as it recovers its normal function (Hudson et al., 2014). As the brain systematically passes through these distinct states of neuronal activity, consciousness may “increase” as the brain recovers, possibly gaining entropy (as postulated by the Entropic Brain Hypothesis). Therefore, understanding how comatose and TBI patients with reduced consciousness recover may provide insight into the levels of human consciousness affected by varied brain states and regions.

Brain Computer Interfaces (BCIs) may even be able to preserve minimally conscious states by following commands and executing external behaviors (Galiotta et al., 2023). Doctors can diagnose and treat disorders of consciousness and locked-in syndrome by utilizing BCIs, which are external devices that detect mental activities like motor imagery, spatial navigation, and arithmetic to allow patients to engage with the world. Advances such as these push us to question how much a physical body, which can be plagued by disease, is needed to interact with the world (Manisha et al., 2022).

"Doctors can diagnose and treat disorders of consciousness and locked-in syndrome by utilizing BCIs, which are external devices that detect mental activities like motor imagery, spatial navigation, and arithmetic to allow patients to engage with the world"

Conclusion

Characterizing consciousness is the fundamental basis of many important questions, whether it be the discontinuity of life support for coma patients, personality shift following brain injury, or even the consciousness of brain organoids and AI models. By theorizing and measuring consciousness, we can understand how treatment affects neurological disorders. Therapeutic pathways ranging from meditation to immune pathway harnessing have an effect on consciousness (Kraemer et al., 2022; Valiukas et al., 2022). Theories, spanning AST to IIT, and tools, like rsfMRI and EEG, may not only provide insight into the effectiveness of treatment, but possibly the identity of the patient themselves. Defining consciousness is no longer just a philosophical debate, but a necessary area of study to cure the ailments of the human brain.

References

Bodhi, Bhikkhu (Trans.). (2000). The connected discourses of the Buddha: A new translation of the Samyutta Nikāya. Wisdom Publications.

Biringer, F., & Anderson, J. R. (1992). Self-recognition in Alzheimer's disease: A mirror and video study. *Journal of Gerontology, 47*(6), P385–P388. https://doi.org/10.1093/geronj/47.6.p385

Carhart-Harris, R. L. (2014). The entropic brain: A theory of conscious states informed by neuroimaging research with psychedelic drugs. Frontiers in Human Neuroscience, 8, 20. https://doi.org/10.3389/fnhum.2014.00020

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., Casarotto, S., Bruno, M.-A., Laureys, S., Tononi, G., & Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine, 5*(198), 198ra105. https://doi.org/10.1126/scitranslmed.3006294

Dehaene, S. (2016). Consciousness and the brain: Deciphering how the brain codes our thoughts. *Neuroscience & Biobehavioral Reviews, 83*, 229–242. https://doi.org/10.1016/j.neubiorev.2016.01.002

Dehaene, S., Charles, L., King, J. R., & Marti, S. (2011). Neural correlates of consciousness: Progress and problems. *Neuroscience & Biobehavioral Reviews, 36*(2), 775–792. https://doi.org/10.1016/j.neubiorev.2011.07.001

Eliasmith, C. (n.d.). Semantic Pointer Architecture [Slides]. Slides by Xueyang Yao. https://ndey96.github.io/deep-learning-paper-club/slides/Semantic%20Pointer%20ArchiEcture.pdf

Galiotta, V., Luo, L., Haugg, A., Kotchoubey, B., & Naci, L. (2023). EEG-based brain–computer interfaces for people with disorders of consciousness: Features and applications. A systematic review. *Frontiers in Human Neuroscience, 17*, 1117653. https://doi.org/10.3389/fnhum.2023.1117653

Graziano, M. S. A., & Webb, T. W. (2015). The attention schema theory: A mechanistic account of subjective awareness. *Frontiers in Psychology, 6*, 500. https://doi.org/10.3389/fpsyg.2015.00500

Hudson, A. E., Calderon, D. P., Pfaff, D. W., & Proekt, A. (2014). Recovery of consciousness is mediated by a network of discrete metastable activity states. *Proceedings of the National Academy of Sciences of the United States of America, 111*(25), 9283–9288. https://doi.org/10.1073/pnas.1408296111

Kant, I. (1998). Critique of pure reason (P. Guyer & A. W. Wood, Trans.). *Cambridge University Press*. (Original work published 1787)

Kraemer, K. M., Jain, F. A., Mehta, D. H., & Fricchione, G. L. (2022). Meditative and mindfulness-focused interventions in neurology: Principles, science, and patient selection. *Seminars in Neurology, 42*(2), 123–135. https://doi.org/10.1055/s-0042-1742287

Locke, J. (1689). An essay concerning human understanding. Thomas Basset.

Manisha, M., Raza, H., Cecotti, H., Prasad, G., & Majid, A. (2022). EEG-based brain-computer interfaces for communication and rehabilitation of people with motor impairment: A novel approach of the 21st century. *Frontiers in Neuroscience, 16*, 959339. https://doi.org/10.3389/fnins.2022.959339

Nagel, T. (1974). What is it like to be a bat? The Philosophical Review, 83(4), 435–450.

Richle, M. E., & Mintun, M. A. (2006). Brain work and brain imaging. *Annual Review of Neuroscience, 29*, 449–476. https://doi.org/10.1146/annurev.neuro.29.051605.112819

Seth, A. K., & Bayne, T. (2021). Theories of consciousness. *Nature Reviews Neuroscience, 22*(8), 514–526. https://doi.org/10.1038/s41583-022-00587-4

Simic, G., Babic, M., Borovecki, F., & Hof, P. R. (2014). Early failure of the default-mode network and the pathogenesis of Alzheimer's disease. *CNS Neuroscience & Therapeutics, 20*(7), 692–698. https://doi.org/10.1111/cns.12260

Smith, D. W. (2013). Phenomenology. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2013 Edition). Stanford University. https://plato.stanford.edu/entries/phenomenology/

Tye, M. (2021). Qualia. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2021 Edition). Stanford University. https://plato.stanford.edu/entries/qualia/

Valiukas, Z., Ephraim, R., Tangalakis, K., Davidson, M., Apostolopoulos, V., & Feehan, J. (2022). Immunotherapies for Alzheimer's disease—A review. *Vaccines, 10*(9), 1527. https://doi.org/10.3390/vaccines10091527

Van Gulick, R. (2023). Consciousness. In E. N. Zalta & U. Nodelman (Eds.), The Stanford Encyclopedia of Philosophy (Winter 2023 Edition). Stanford University. https://plato.stanford.edu/entries/consciousness/