JNphi

Neurophilosophy of Consciousness: From Biological Basis to Subjective Reality

Gonzalo Emiliano Aranda-Abreu

Abstract

Motivated by the persistent challenges posed by linking neural dynamics with subjective experience, and based on previous philosophical and neuroscientific research, this paper explores consciousness as an emergent, complex, and multidimensional property rooted in biological systems and shaped by evolutionary imperatives of survival. Based on a critical integration of neuroscience, philosophy, and cognitive science, a unified theory is articulated that links the neurobiological foundations of the conscious state, such as the role of the brain stem, thalamus, synapses, and neurotransmitters, with computational models of predictive processing, global diffusion, and information integration. It is argued that consciousness is not a phenomenon of all or no but rather emerges gradually from homeostatic affective levels toward more complex forms of self-awareness and metacognition. Furthermore, consciousness is analyzed as an active simulation of the environment, challenging naive realism and reframing the problem of free will from the perspective of predictive processing. A five-level hierarchical model is proposed, ranging from homeostatic regulation to metacognition and phenomenological complexity. Finally, ethical and philosophical implications arising from advances in neuroscience and biotechnology, such as brain organoids, are addressed. The aim is of consciousness, but to transform it into a series of questions that can be addressed through an integrated scientific and philosophical approach.

Key Words: Consciousness; Neurophilosophy; Predictive Processing; Computational Integration; Subjective Reality

DOI: 10.5281/zenodo.20051601

Corresponding author: Gonzalo Emiliano Aranda-Abreu
Address: Instituto de Investigaciones Cerebrales/Universidad Veracruzana, Xalapa, Ver. Mexico.
e-mail ✉ garanda@uv.mx

1. Introduction

Rene Descartes, in his so-called Cartesian dualism, and the division between extended matter (*res extensa*) and thinking substance (*res cogitans*), established the historical framework for the mind-body problem (Camargo, 2025). For centuries, this separation has shaped the debate on the nature of the mind. This challenge is persistent and encapsulated in what philosopher David Chalmers called the “hard problem” (Chalmers, 1995) of consciousness: the argument as to why processing is accompanied by a subjective first-person experience.

This dilemma, articulated by Joseph Levine as an explanatory gap, highlights the apparent impossibility of deducing the phenomenal quality of experience, *qualia*, such as the redness of red or the sweet taste of honey, from a purely physical description of neural activity. (Levine, 1983) Subjective experience resists being reduced to mere nerve impulses and leaves a conceptual void that no yet can be explained in scientific terms (Schiffer, 2019).

The aim of this manuscript is to articulate a unified theory of consciousness that overcomes this traditional separation, integrating findings from neuroscience with conceptual models from philosophy and cognitive science. It seeks a single “essence” of consciousness, proposing that this is an emergent property of a complex biological system. For this construction, it is first necessary to analyze the biological foundations that make the conscious state possible (Figure 1).

For conceptual clarity, it is useful to distinguish between different forms of consciousness that are addressed throughout this manuscript. Arousal refers to the physiological state of wakefulness; phenomenal consciousness to subjective experience or “what-it-is-like”; access consciousness to the availability of information for reasoning, report, and control; self-consciousness to representations of oneself as an experiencing subject; and metacognition to higher-order awareness of mental states. While these dimensions are deeply interrelated, they are not identical. The present approach aims to account for their interaction without reducing them to a single undifferentiated concept, in line with the distinctions proposed by philosophers such as Ned Block (Block, 2007).

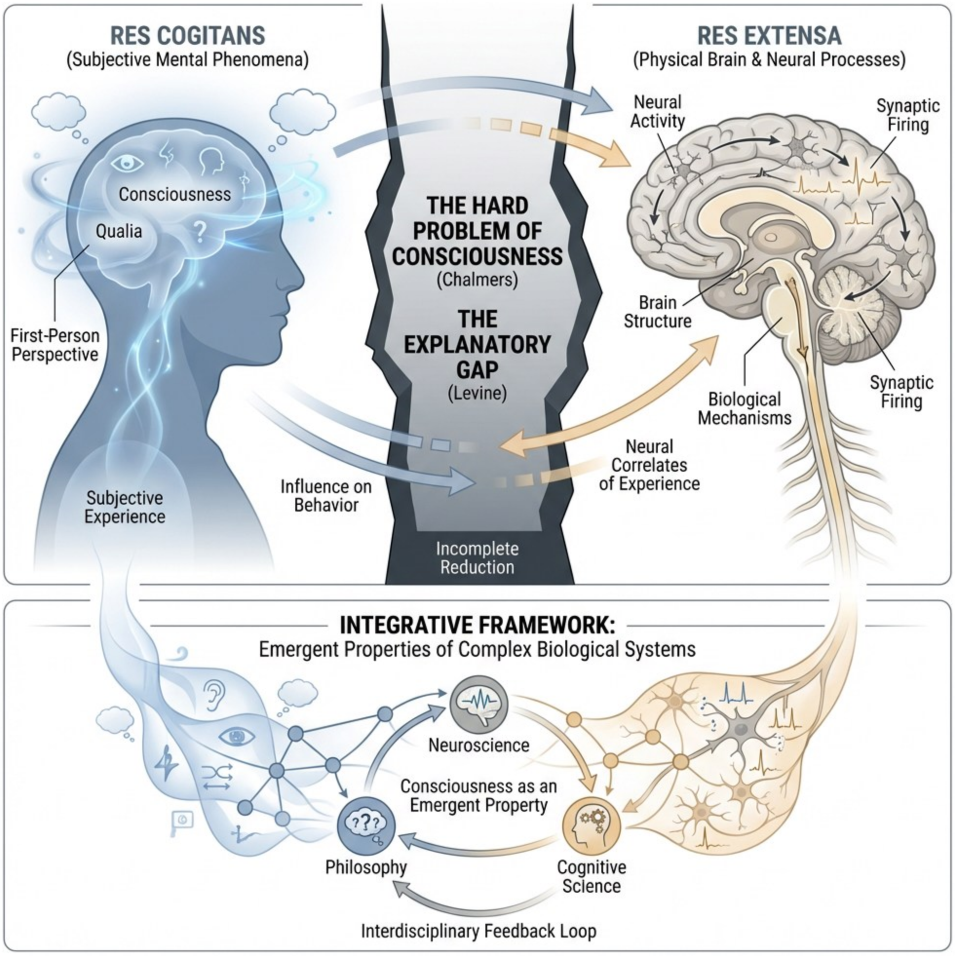

Figure 1. Conceptual schematic of the mind–body problem and its contemporary reformulation. The figure contrasts res cogitans, representing subjective mental phenomena such as consciousness, qualia, first-person perspective, and lived experience, with res extensa, representing the physical brain, including neural activity, synaptic firing, brain structure, and underlying biological mechanisms. The central fissure symbolizes the "hard problem of consciousness" (Chalmers) and the explanatory gap (Levine), highlighting the persistent difficulty of deriving subjective experience from purely physical descriptions of neural processes and the incomplete reduction of mind to brain activity. Bidirectional arrows indicate correlations and causal influences between neural mechanisms and behavior without implying full explanatory closure. The lower panel introduces an integrative framework in which consciousness is conceptualized as an emergent property of complex biological systems, arising from dynamic interactions studied across neuroscience, philosophy, and cognitive science, and sustained through interdisciplinary feedback loops that aim to bridge, rather than eliminate, the mind–brain divide.

2. Neurobiological Foundations of Consciousness

The intricate nature of subjective experience is an essential component for understanding the neurobiological processes that enable the state of wakefulness, which is a *sine qua non* condition for consciousness to manifest itself in its everyday form (Alcaraz-Sánchez, 2024). Consciousness does not navigate freely; it is anchored to a complex biological system whose functioning is indispensable. This analysis of the biological "element" provides the empirical anchor necessary for more abstract theories, ensuring that any model of consciousness remains connected to the physiological reality of the brain.

2.1 The Role of the Brain Stem and Thalamus in Cortical Activation

The study of the mechanisms of wakefulness began with experiments conducted by Bremer in the 1930s. (Kerkhofs and Lavie, 2000) His studies on cats, in which he performed the "*cerveau isolé*" preparation, consisting of dissecting the midbrain between the inferior and superior colliculi, induced a state similar to deep sleep, suggesting that the brain requires ascending afferents to remain awake. Subsequent studies identified the critical structures more precisely. Fuller et al. (2011) demonstrated that intracerebral injection of saporin into the medial parabrachial nucleus of rats induced a persistent comatose state, confirming the critical role of this structure in wakefulness. EEG recordings showed continuous delta activity, and recovery of consciousness was only possible by stimulating brainstem nuclei with precision optogenetics (Fuller et al., 2011). These findings converge with human studies: analyses of brain lesions causing coma, such as those by Fischer et al., demonstrated that focal damage to the pontine tegmentum (dorsal region of the pons (Varolian bridge), which is part of the brainstem) specifically in or near the medial parabrachial nucleus, is significantly more likely in comatose patients (Fischer et al., 2016).

The thalamus, for its part, acts as a crucial bridge between the brain stem and the cortex. Although studies such as those by Fuller et al. showed that extensive thalamic lesions no eliminate the sleep-wake cycle, more recent research has revealed that specific nuclei, particularly the intralaminar nuclei, promote cortical activation (Honjoh et al., 2018). Therefore, the brainstem appears to contain the neurons essential for wakefulness, while the thalamus plays a key modulatory role in orchestrating the cortical activity that characterizes the conscious state.

2.2 Synaptic and Molecular Basis of Neuronal Activity

At the molecular level, the state of consciousness is modulated by neurotransmitters, receptors, and synaptic processes (Teleanu et al., 2022). Anesthetics, for instance, induce loss of consciousness by enhancing the action of the neurotransmitter GABA, acting on GABA-A receptors to suppress widespread neuronal activity (Son, 2010).On the excitatory side, NMDA receptors—crucial for synaptic plasticity such as long-term potentiation (LTP)—are essential for the maintenance of consciousness (Sumi and Harada, 2020). This leads us to emphasize that, without receptors and neurotransmitters, a conscious state would not be possible. As highlighted in the studies conducted by Min Zhuo, long-term potentiation (LTP) is a fundamental mechanism of learning and memory, functions that are intrinsically linked to conscious experience (Zhuo, 2024). The importance of these receptors is underscored in cases of anti-NMDA receptor encephalitis, where the patient's immune system attacks these receptors, leading to severe cognitive deficits and loss of consciousness (Nichols, 2016). Other cells such as astrocytes also play a regulatory role, as they secrete proteins like Neurocan, which controls the formation of specific inhibitory synapses (particularly those on somatostatin-expressing interneurons), illustrating the complexity of the cellular network that supports brain function and, by extension, consciousness (Irala et al., 2024).

2.3 The Affective Origin: Homeostasis and Feeling as a Source of Consciousness

Why did consciousness evolve? Antonio Damasio proposes an important hypothesis: consciousness did not emerge as a high-level cognitive capacity, but from the very foundations of life itself. (Damasio and Damasio, 2023) According to Damasio, homeostatic feelings—the subjective experiences of pain, well-being, hunger, thirst, or discomfort—are not merely contents of consciousness, but a primary evolutionary source. These feelings provide direct “knowledge” about the state of life regulation within the organism (Damasio and Damasio, 2022). Knowing whether we are in an optimal or dangerous state offers a fundamental adaptive advantage, allowing us to guide behavior in a much more flexible and effective manner than through automated reflexes alone. In this view, affective consciousness is the first step: the conceptual bridge that connects the biological mechanisms of survival with subjective emergence, laying the groundwork for more complex forms of self-awareness. (Weissman et al., 2020) Once these biological foundations are established, the next step is to examine how neuronal activity is organized into processing systems that generate conscious experience (Figure 2).

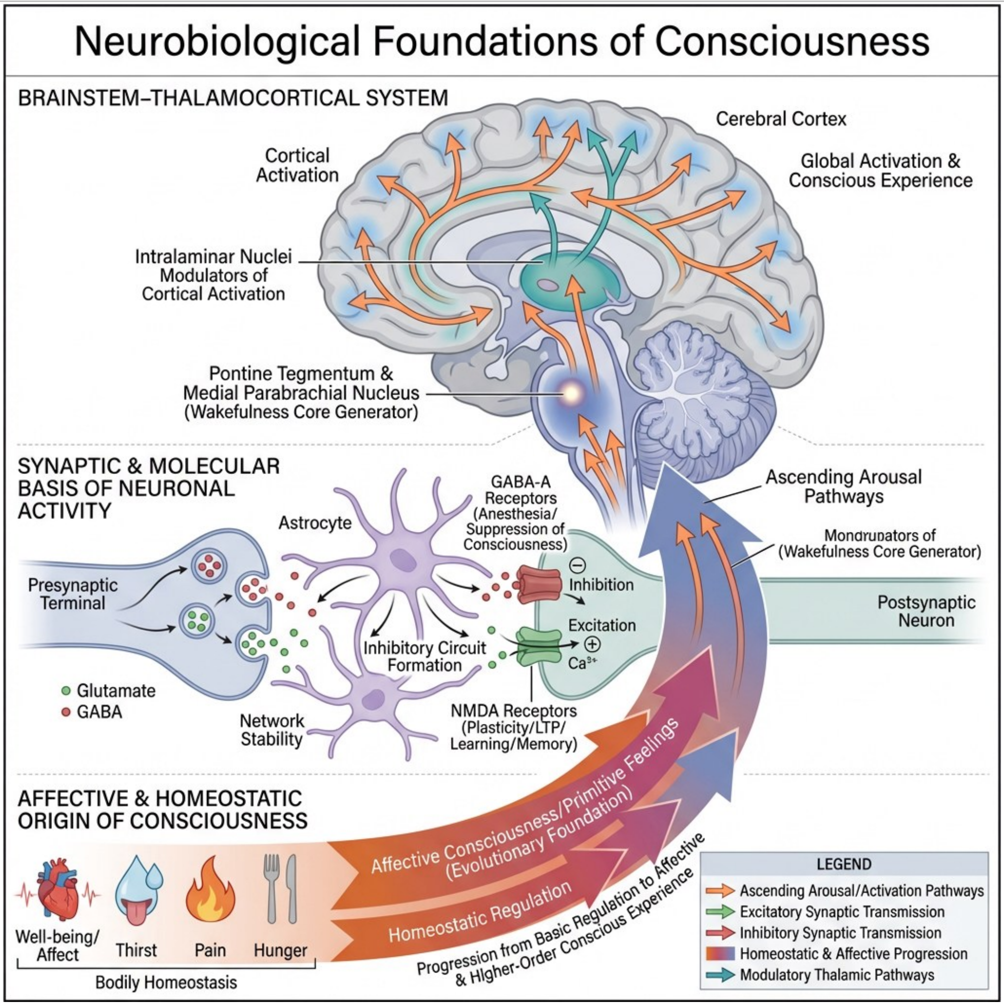

Figure 2. Neurobiological foundations of consciousness. This schematic summarizes the biological mechanisms that enable wakefulness and conscious experience across hierarchical levels of brain organization. The upper panel illustrates the brainstem–thalamocortical system, highlighting the pontine tegmentum and medial parabrachial nucleus as core generators of wakefulness that provide ascending arousal signals to the thalamus and cerebral cortex. Intralaminar thalamic nuclei are depicted as key modulators of cortical activation, coordinating widespread cortical activity necessary for global conscious experience. The middle panel depicts the synaptic and molecular basis of neuronal activity, emphasizing the balance between excitation and inhibition that sustains functional neural networks. Excitatory glutamatergic transmission mediated by NMDA receptors supports synaptic plasticity, long-term potentiation, learning, and memory, whereas inhibitory signaling through GABA-A receptors suppresses neuronal activity and underlies states such as anesthesia and loss of consciousness. Astrocytes are shown as active regulators of synaptic organization and inhibitory circuit formation, contributing to network stability and information processing. The lower panel represents the affective and homeostatic origin of consciousness, in which bodily signals related to pain, thirst, hunger, and well-being generate primitive feelings that provide the evolutionary foundation of conscious experience. These homeostatic and affective processes progressively integrate with ascending arousal systems and cortical networks, giving rise to higher-order forms of consciousness. Together, the figure illustrates consciousness as an emergent phenomenon grounded in coordinated brainstem, thalamic, cortical, synaptic, and affective mechanisms.

124

3. Models of Conscious Processing

Beyond the biological substrates that enable wakefulness, various theoretical frameworks compete to elucidate the functional and computational architecture of consciousness. These proposals aim to clarify how neuronal activity gives rise to a subjective, integrated, and coherent experience. Despite their conceptual divergences, these theories can be grouped into:

(i) Global integration and broadcast models, which emphasize the importance of large-scale brain connectivity; (ii) The predictive processing framework, which conceives the brain as a hierarchical inference-generating system; and (iii) Higher-order theories, which focus on the metacognitive and self-referential dimensions of consciousness.

3.1 Theories of Global Integration

Two of the most influential theories in this domain are the Global Workspace Theory (GWT), originally proposed by Bernard Baars (Baars, 1993) and later developed by Stanislas Dehaene, (Dehaene and Naccache, 2001) and the Integrated Information Theory (IIT) by Giulio Tononi. (Tononi et al., 2016)

GWT conceptualizes consciousness as a phenomenon of “fame in the brain.” According to this model, information becomes conscious when it is selected and broadcast across a “workspace” composed of long-range neurons, primarily within the prefrontal and parietal cortices. This global broadcast—or ignition—makes information accessible to a wide range of cognitive processes, including working memory, planning, and verbal report. In this view, consciousness is synonymous with global access.

In contrast, IIT approaches consciousness from a phenomenological perspective. It posits that consciousness is an intrinsic property of systems capable of integrating information in an irreducible manner. This integrative capacity is quantified by a parameter known as Phi (Φ), which reflects the degree to which a system generates a unified informational whole that exceeds the mere sum of its parts. A system is thus considered conscious to the extent that its structural organization produces information that cannot be decomposed without significant loss—that is, when the whole is greater than the sum of its components.

Both theories converge on the idea that consciousness involves large-scale integration of neuronal activity, but they diverge in focus: GWT emphasizes function and access, while IIT emphasizes the intrinsic structure of experience.

3.2 The Predictive Brain: Consciousness as Active Inference

The predictive processing (PP) paradigm, popularized by theorists such as Andy Clark (Clark, 2016) and Anil Seth, (Seth, 2021) offers a fundamentally different view of brain function. According to this model, perception is not a passive, bottom-up process in which the brain merely receives and assembles sensory data. Instead, it is an active, top-down process in which the brain constantly generates hypotheses or predictions about the causes of its sensory inputs. As Daniel Yon (Yon, 2026) puts it, perception is a “controlled hallucination”. The brain builds a model of the world and continuously updates it based on prediction error—the difference between what it expects and what the senses actually report.

In functional magnetic resonance imaging (fMRI) studies, Kok et al. (2012) demonstrated that visual predictions generated by the prefrontal cortex modulate activity in V1 even before stimulus presentation, suggesting an active top-down inference process (Kok et al., 2012). Furthermore, prediction errors have been associated with gamma-band desynchronization detected via MEG recordings, (Bastos et al., 2015) providing neurophysiological support for Seth’s “controlled hallucination” hypothesis.

This top-down model finds its neurobiological counterpart in Recurrent Processing Theory (RPT), proposed by Victor Lamme, (Lamme, 2006) which posits that recurrent feedback loops between high- and low-level cortical areas are essential for sensory information to become conscious. Initial feedforward processing remains unconscious. Conscious experience emerges only when high-level predictions recurrently modulate lower-level sensory processing through these feedback loops.

3.3 Higher-Order Representations and Self-Consciousness

Higher-order theories (HOT), developed by philosophers such as David Rosenthal, (Rosenthal, 2008) directly address the reflexive aspect of consciousness: the fact that we not only have experiences, but we are also aware of having them. According to this view, a mental state (such as a perception or emotion) is not intrinsically conscious. It becomes conscious only when it becomes the object of a meta-mental state—specifically, a “higher-order thought” about it. Being aware of a pain is not just having the pain state, but having a thought of the form: *“I am in a state of pain.”*

This approach contrasts with first-order theories such as GWT and IIT, which focus on the intrinsic properties of the mental state itself—e.g., its global broadcast or its level of integrated information. HOT theories, in contrast, explain consciousness as a relation between mental states, emphasizing the role of metacognition and the capacity for self-representation.

These theoretical models aim to explain the mechanisms that shape the intrinsically constructed nature of our subjective reality (Figure 3).

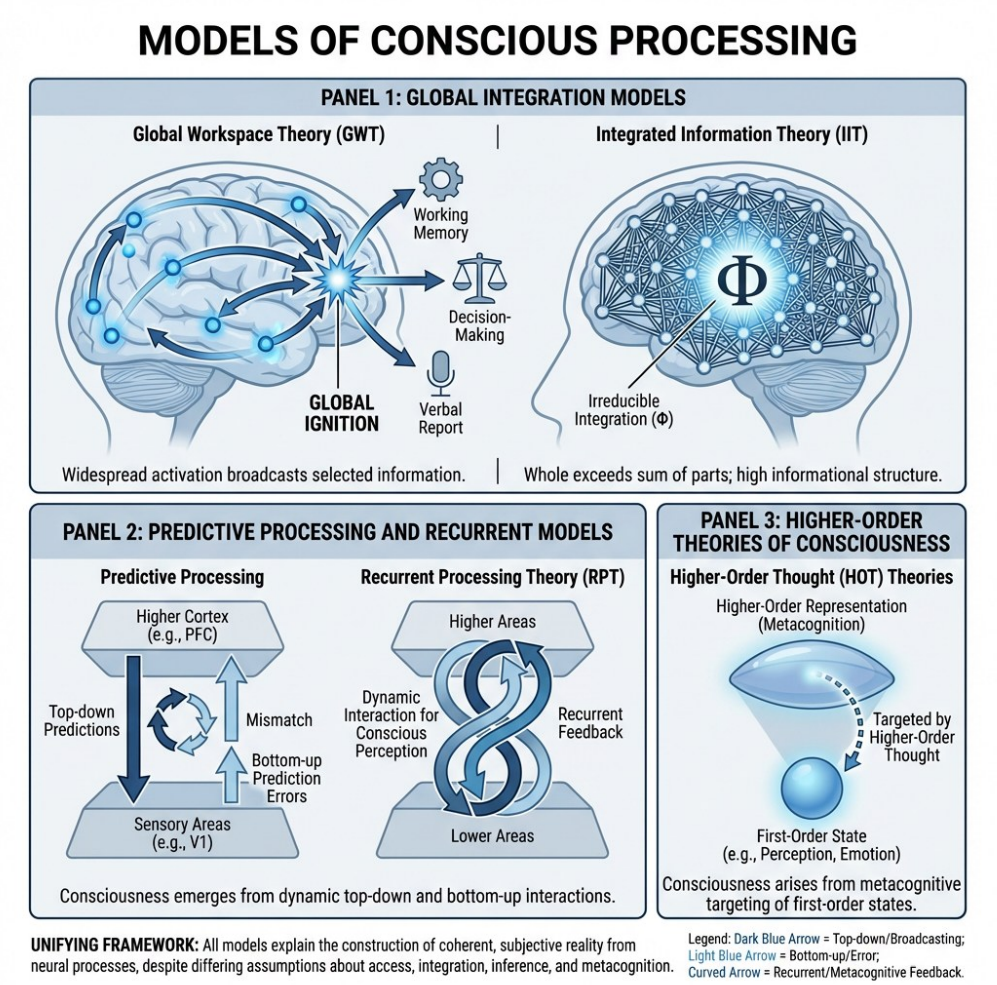

Figure 3. Models of conscious processing. This figure summarizes major theoretical frameworks that seek to explain how neural activity gives rise to coherent, subjective conscious experience. Panel 1 presents global integration models, contrasting Global Workspace Theory (GWT)—in which selected information becomes conscious through global ignition and widespread broadcasting across frontoparietal networks, enabling working memory, decision-making, and verbal report—with Integrated Information Theory (IIT), which conceptualizes consciousness as an intrinsic property of systems capable of irreducible information integration, quantified by Φ (phi), where the whole exceeds the sum of its parts. Panel 2 illustrates predictive processing and recurrent models. In the predictive processing framework, higher cortical areas generate top-down predictions that shape sensory processing, while bottom-up prediction errors signal mismatches between expectations and sensory input. Recurrent Processing Theory (RPT) emphasizes dynamic feedback loops between higher and lower cortical areas, proposing that conscious perception emerges only when recurrent interactions supplement initial feedforward processing. Panel 3 depicts higher-order theories of consciousness, according to which a first-order mental state (such as a perception or emotion) becomes conscious only when it is the target of a higher-order, metacognitive representation, highlighting the reflexive and self-referential dimension of conscious awareness.

4. The Nature of Subjective Reality: The World as a Neural Construction

The processing models described above converge on a radical and counterintuitive conclusion: our experience of the world, of the self, and of agency is not a direct or passive reflection of external reality, but a sophisticated and active construction by the brain. Far from being a transparent window onto the world, consciousness is more like to a simulation model that the brain generates in order to effectively navigate and interact with the environment.

4.1 The Grand Illusion: Deconstructing Perception

The belief that we perceive the world as it truly is—a stance known as naïve realism—is a persistent illusion. Daniel Dennett (Dennett, 1991) coined the term *"Cartesian theater"* to critique the idea that there exists a central location in the brain where "everything comes together" and is presented to an internal observer. Phenomena such as change blindness and inattentional blindness, demonstrated in the studies of Susan Blackmore (Blackmore, 2002) and Ronald Rensink, (Rensink, 1997) reveal that our perception of a rich and complete visual world is largely a fabrication. We do not process every detail of our environment; rather, the brain captures the "gist" or summary and fills in the rest based on its own expectations.

As Charles Pinter (Pinter, 2020) argues, even the segmentation of the world into discrete objects and the organization of scenes according to Gestalt principles (such as proximity, similarity, and closure) are not inherent properties of matter, but structures actively imposed by the mind to make sense of sensory input.

4.2 The Construction of the "Self" and the Problem of Free Will

Just like our perception of the external world, the experience of a unified and persistent "self" is also a neural construction. The philosopher David Hume, in his bundle theory, argued that the self is not a substantial entity, but simply a collection of flowing perceptions (Hume, 2020). This idea resonates in modern neuroscience with Michael Gazzaniga's concept of the *"left-hemisphere interpreter."* Based on split-brain patient studies, Gazzaniga proposed that the left hemisphere constantly generates a cohesive narrative to make sense of our actions and feelings, functioning as a *"narrative center of gravity"* that creates the illusion of a coherent self. (Gazzaniga, 1992)

This view of the self as a *post-hoc* construct challenges the traditional notion of free will. The experiments of Benjamin Libet (Libet et al., 1983) demonstrated that brain activity related to voluntary action precedes the conscious experience of the decision to act. Building on this, psychologist Daniel Wegner (Wegner, 2003) proposed the *"illusion* *of conscious will"*, arguing that the experience of agency is not the direct cause of our actions. Instead, we infer that our thoughts caused an action if certain criteria are met (such as priority and consistency), but both the thought and the action are actually products of unconscious brain processes.

From the perspective of predictive processing, the experience of agency is merely a postdictive inference by the brain: it infers that its own predictions about action matched the sensory feedback, thus generating a model of the *"self"* as the most probable cause. While classical experiments by Libet and subsequent interpretations by Wegner have played a pivotal role in questioning traditional notions of free will, their conclusions remain the subject of substantial debate. Methodological critiques have highlighted limitations related to temporal resolution, task artificiality, and the interpretation of readiness potentials. Within the present framework, the postdictive model of agency is therefore not treated as an absolute denial of free will, but as a probabilistic account of how the experience of agency is constructed. Conscious intention is understood as one component within a broader predictive and inferential process, rather than as a singular causal origin of action.

4.3 Animal Consciousness and the Degrees of Experience

If consciousness is a product of biological evolution, it should not be an all-or-nothing phenomenon exclusive to humans. A multidimensional framework, as proposed by Jonathan Birch, Alexandra Schnell, and Nicola Clayton, (Birch et al., 2020) suggests that consciousness varies along a spectrum and across multiple dimensions. They propose five key axes for assessing animal consciousness:

1. Perceptual richness: The complexity and detail of sensory experience.
2. Evaluative richness: The capacity to experience a range of affective states (such as pleasure, pain, fear).
3. Unity: The degree to which experience is integrated into a coherent whole at a given moment.
4. Temporality: The ability to integrate experiences over time, such as episodic memory.
5. Self-awareness: The ability to recognize oneself as a distinct individual.

Empirical evidence supports this graded view. The mirror test, which evaluates a form of self-awareness, has been passed not only by great apes but also by dolphins, Asian elephants, magpies, and even the cleaner wrasse fish. These findings suggest a deep evolutionary origin and a graded nature of subjective experience, challenging a purely anthropocentric perspective (Figure 4).

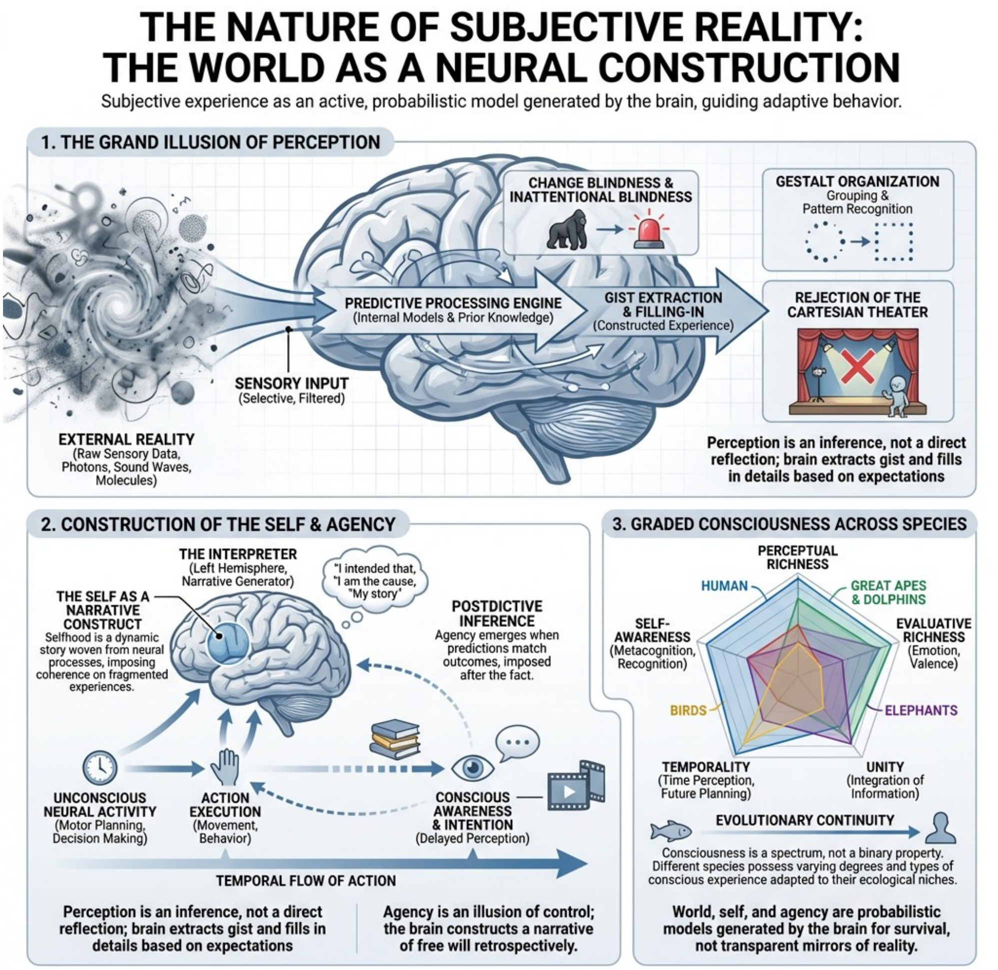

Figure 4. The nature of subjective reality as a neural construction. This schematic illustrates the view that subjective experience is not a direct reflection of external reality, but an active, probabilistic model generated by the brain. (1) The grand illusion of perception: raw sensory inputs from the external world are selectively filtered and processed through predictive mechanisms, whereby the brain extracts the “gist” of sensory information and fills in missing details based on prior knowledge and expectations. Phenomena such as change blindness, inattentional blindness, Gestalt organization, and the rejection of a centralized “Cartesian theater” highlight perception as an inferential process rather than a transparent mirror of the world. (2) Construction of the self and agency: the sense of a unified self and voluntary control emerges as a narrative construct, generated post hoc by neural processes. Unconscious neural activity precedes action execution, while conscious awareness and intention arise later, giving rise to a postdictive inference of agency and the experiential illusion of free will. (3) Graded consciousness across species: consciousness is depicted as a multidimensional and evolutionary continuum, varying across perceptual richness, evaluative richness, unity, temporality, and self-awareness among different species. Together, the figure emphasizes that world, self, and agency are adaptive neural constructions shaped by predictive processing and evolutionary pressures, rather than faithful representations of objective reality.

If consciousness is such a complex and constructed biological phenomenon, how might the various theories converge to form a coherent framework?

5. Toward a Unified Theory

Despite their apparent differences, contemporary theories of consciousness share significant points of convergence that allow for the sketching of a unified, multi-scale framework. This integrative approach does not seek a "winning theory," but rather acknowledges that different models may be describing complementary aspects of a coherent and multifaceted phenomenon, highlighting both promising insights and speculative frontiers.

At this stage, it is important to clarify the precise sense in which the present framework is described as a "unified theory" of consciousness. The unification proposed here is not ontological in the strong metaphysical sense, nor does it claim to identify a single fundamental essence of consciousness. Rather, the framework advances an epistemological and explanatory integration across multiple levels of analysis. It aims to coherently map complementary mechanisms—biological, computational, cellular, network-level, and phenomenological—that together constrain and shape conscious experience. In this sense, unity is achieved through cross-level coherence and mutual constraint, rather than through reductive explanation or metaphysical identity.

5.1 Multiple Scales: Recurrence, Integration, and Prediction

A synthetic model, based on convergences identified in integrative works such as that of Storm et al. (Storm et al., 2024), can be structured across the following hierarchical levels:

1. Homeostatic Foundation: At its most basic level, consciousness is anchored in the biological imperative of life regulation. As proposed by Antonio Damasio, affective feelings (pleasure, pain, well-being) act as the primordial evolutionary driver, providing the organism with first-person knowledge of its internal state and guiding adaptive behavior. (Damasio and Damasio, 2023)
2. Predictive Mechanism: At the computational level, the brain operates as a Bayesian inference machine. Through the principles of Predictive Processing (PP) and Recurrent Processing Theory (RPT), it continuously generates a model of the world and of itself in order to minimize prediction error. Conscious experience results from this ongoing process of hypothesis generation and revision.
3. Dendritic Integration: At the cellular level, the integration of feedback and sensory input signals occurs in the apical

dendrites of cortical neurons. This mechanism, emphasized by Aru and Larkum**,** (Aru et al., 2012) functions as a fundamental computational unit for conscious processing, allowing context and experience to modulate sensory perception.

4. Global Broadcast and Metacognition: When a prediction error is particularly salient or the information is highly relevant to the organism's goals, an "ignition" event occurs. This central concept of Global Workspace Theory (GWT) involves the broadcasting of information through a fronto-parietal network, making it accessible to high-level cognitive control, planning, and verbal report. This process is intrinsically linked to metacognition. As suggested by Chris Frith**,** (Shea and Frith, 2019) the brain assigns a confidence level to information, creating a "common currency" that allows data from different modalities (perceptual, mnemonic, affective) to be compared and integrated for decision-making.
5. Complexity and Differentiation: Phenomenologically, conscious states are characterized by high complexity, defined in Integrated Information Theory (IIT) as the coexistence of high integration (a unified experience) and high differentiation (rich in distinct contents). Rather than equating consciousness with high entropy per se, a more precise characterization is required. Empirical evidence suggests that conscious states are associated with a broad and flexible dynamical repertoire, characterized by a balance between integration and variability (Jang et al., 2024). Deep anesthesia is associated with reduced dynamical complexity, psychedelic states with expanded variability, and epileptic seizures with pathological hyper-synchrony despite high signal power (Palencia, 2025). Accordingly, the present framework refers not to entropy alone, but to flexible state dynamics or dynamic repertoire breadth, reflecting the system's capacity to explore a rich yet constrained space of neural configurations compatible with integrated experience.

A further conceptual distinction must be emphasized between neural correlates of consciousness (NCCs) and genuinely causal mechanisms. Within the present framework, not all identified neural processes are proposed as sufficient causes of consciousness. Rather, they play different explanatory roles(Mckilliam, 2024). Brainstem arousal systems and thalamocortical activation are treated as necessary enabling conditions; dendritic integration and recurrent processing function as facilitating mechanisms that shape conscious content; while global workspace ignition and metacognitive access operate as access-amplifying and report-enabling processes(Bast et al., 2025). Conscious experience is therefore not localized to a single causal locus, but emerges from the coordinated interaction of multiple necessary and facilitating components, none of which alone is sufficient.

The apparent compatibility between Predictive Processing (PP) and Integrated Information Theory (IIT) warrants careful qualification. These frameworks are grounded in distinct ontological commitments: IIT characterizes consciousness as an intrinsic property of certain physical systems, whereas PP interprets conscious experience as the functional outcome of hierarchical inference and error minimization. In the present model, these perspectives are not treated as interchangeable, but as addressing different explanatory questions. PP primarily accounts for the dynamic construction and content of experience, while IIT provides constraints on the structural conditions under which unified experience is possible. Their coexistence within this framework is therefore pluralistic rather than reductive, acknowledging theoretical tension while preserving complementary explanatory value(Bartlett, 2022).

5.2 Challenges and Future Perspectives: From Quantum Physics to Neuroethics

Epistemological difficulty is not an obstacle to avoid, but a territory to explore with intellectual honesty. Integrating it into scientific and philosophical discourse on consciousness is a prerequisite for a truly comprehensive, open, and critical theory. A philosophical position known as mysterianism (Colin McGinn) suggests that consciousness may be cognitively irresolvable for humans—just as a bat cannot grasp Euclidean geometry. While debatable, this stance offers a healthy epistemic caution: the study of consciousness must advance without losing sight of its limits or reifying an experience that, by definition, resists purely physical explanations. (Cavanna and Nani, 2014)

Empirical validation of consciousness is especially ambiguous in non-human contexts. In non-verbal organisms, AI, or brain organoids, what constitutes sufficient evidence? Behavior? Connectivity? Complexity? Without clear criteria, we risk implicit anthropocentrism or projecting our own cognitive structures inappropriately.

Even the most promising theories—such as IIT, PP, and HOT—are informed by implicit ontological and epistemological assumptions. For example, IIT assumes consciousness to be a quantifiable, intrinsic property of physical systems, while HOT theories rely on a representationalist conception of mind. These frameworks are not neutral; they are shaped by prior philosophical commitments that must be acknowledged to evaluate their scope and limitations.

In line with Thomas Nagel's famous essay *"What is it like to be a bat?"*, the subjective nature of conscious experience cannot be fully captured by objective science. We cannot know what it is like to be another conscious being without adopting their first-person perspective. This gap challenges the positivist ideal of a complete science of mind.(Nagel, 1980)

A major epistemological challenge is the distinction between neural correlates of consciousness (NCCs) and causal mechanisms. Many studies have identified brain patterns associated with conscious states (e.g., fronto-parietal activation, recurrent loops, informational complexity), but they do not explain why or how these patterns generate phenomenal experience. This disconnect is what Joseph Levine called the "explanatory gap"—the impossibility of deriving *qualia* (the redness of red, the pain of pain) from a purely functional-physical description.

Despite the significant advances in neuroscience, philosophy of mind, and AI, a central challenge remains: the epistemologically resistant nature of consciousness. This resistance is not merely technical, but conceptual, and reflects a fundamental fissure between scientific language and subjective experience.

5.3 Epistemological Limits in the Study of Consciousness

The boundaries of consciousness research venture into speculative domains and raise profound ethical dilemmas. Theories such as Penrose-Hameroff's, (Hameroff and Penrose, 2014) which postulate quantum processes in neuronal microtubules, or Vitiello's model based on quantum field theory**,** (Vitiello, 2020) attempt to address non-computational aspects of consciousness. While highly speculative and lacking empirical confirmation, they represent efforts to bridge fundamental physics and subjective experience.

More urgently, biotechnological advances pose unavoidable ethical challenges. The ability to create brain organoids in the lab that exhibit complex neuronal activity raises the theoretical possibility of generating "islands of consciousness" disconnected from any body or environment. This forces us to confront their moral and legal status:

How should such entities be treated? Would it be permissible to use them for research? These questions demonstrate that consciousness research is not merely an academic pursuit—it has profound philosophical and social implications that demand urgent debate (Figure 5).

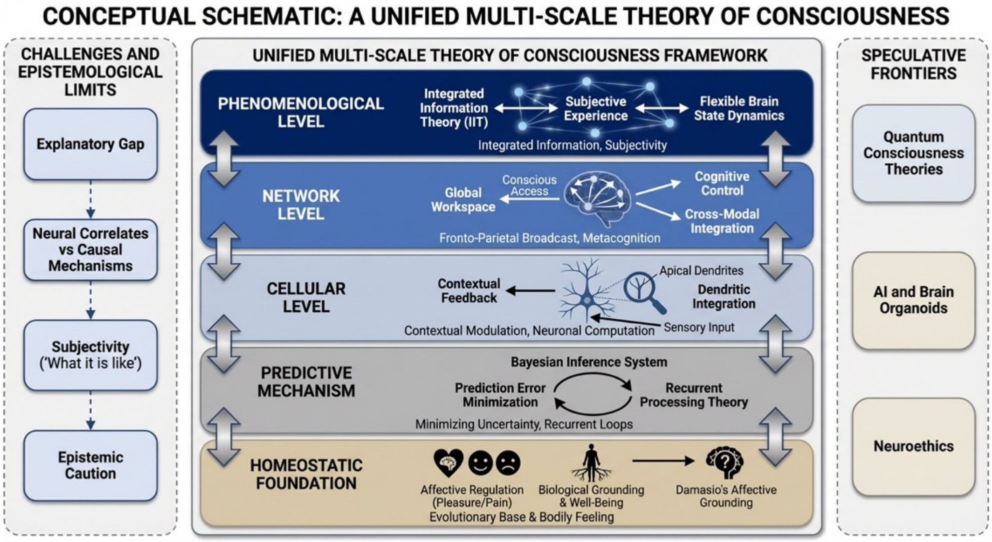

Figure 5. Toward a unified multi-scale theory of consciousness. This conceptual schematic integrates convergent insights from contemporary theories of consciousness into a hierarchical, multi-level framework. At the base, a homeostatic foundation grounds consciousness in biological regulation and affective feelings (pleasure, pain, well-being), emphasizing their evolutionary role in guiding adaptive behavior. Above this, a predictive mechanism depicts the brain as a Bayesian inference system that minimizes prediction error through recurrent processing and top-down predictions. At the cellular level, contextual feedback and sensory input are integrated within apical dendrites, providing a cellular substrate for context-dependent conscious processing. The network level highlights global workspace dynamics, where salient information is broadcast across fronto-parietal networks, enabling conscious access, cognitive control, cross-modal integration, and metacognitive confidence estimation. At the phenomenological level, conscious experience is characterized by high integration and high differentiation, consistent with Integrated Information Theory, and associated with flexible brain state dynamics and increased entropy. The framework is framed by epistemological challenges, including the explanatory gap, the distinction between neural correlates and causal mechanisms, and the irreducibility of first-person subjectivity, as well as by speculative frontiers—such as quantum consciousness theories, artificial intelligence, brain organoids, and neuroethics—highlighting open questions and ethical implications rather than established mechanisms.

6. Conclusion

Rather than seeking finality, our aim is to contribute to a growing interdisciplinary dialogue, one that has guided our own research and philosophical reflections over recent years.

Far from being a “ghost in the machine” or an ineffable quality beyond scientific scrutiny, consciousness emerges as a complex, multidimensional property of biological systems, deeply rooted in the evolutionary imperatives of survival. The integrative perspective presented in this manuscript conceives of consciousness as a dynamic process through which the organism, guided by homeostatic imperatives, actively constructs a coherent predictive model of itself and its environment. This internal model facilitates flexible adaptation, enables goal-directed behavior, and optimizes interaction with a complex and ever-changing world.

The convergence between neuroscience and philosophy does not aim to definitively solve the mystery of consciousness, but rather to reformulate it in tractable terms. By deconstructing this phenomenon into its biological, computational, and phenomenological dimensions, what was once considered an intractable problem is transformed into a structured set of scientific and conceptual questions. In doing so, it generates a fertile framework for exploring one of the most fundamental and defining properties of the human condition.

In summary, the present work offers three principal contributions. First, it advances a biologically grounded account of consciousness that integrates homeostatic, affective, and arousal mechanisms with contemporary computational models. Second, it proposes a multi-scale explanatory framework that reconciles predictive, integrative, and metacognitive approaches without reducing them to a single theoretical axis. Third, it situates consciousness within its epistemological limits, emphasizing the distinction between correlates and causes, and highlighting ethical and conceptual challenges posed by emerging technologies. A central unresolved question remains: how can first-person subjective experience arise from, yet remain irreducible to, third-person biological description? Addressing this question will require not only empirical refinement but sustained philosophical engagement.

Table1. Key Studies on Consciousness

Author(s)	Study / Theory	Model / System	Key Findings
(Kerkhofs and Lavie, 2000)	*Cerveau isolé* cat experiments	Brainstem	Brain requires ascending input to maintain wakefulness.
(Fuller et al., 2011)	Saporin lesion in parabrachial nucleus	Brainstem	Lesions cause coma; recovery via optogenetic stimulation.
(Fischer et al., 2016)	Lesion mapping in comatose patients	Pontine tegmentum	Damage to medial parabrachial nucleus correlated with persistent coma.
(Honjoh et al., 2018)	Thalamic nuclei stimulation	Intralaminar thalamus	Specific nuclei modulate cortical activation and support consciousness.
(Teleanu et al., 2022)	Neurochemical modulation	GABA & NMDA receptors	Anesthetics enhance GABA; NMDA critical for consciousness.
(Sumi and Harada, 2020)	NMDA receptor function	Synaptic plasticity	NMDA receptors support long-term potentiation and awareness.
(Zhuo, 2024)	LTP mechanisms	Memory & Learning	LTP underlies conscious memory and learning

			processes.
(Nichols, 2016)	NMDA receptor encephalitis	Clinical neurology	Immune attack on NMDA receptors causes consciousness loss.
(Irala et al., 2024)	Astrocyte-mediated synapse regulation	Astrocytic proteins	Neurocan regulates inhibitory synapse formation and contributes to awareness.
(Damasio and Damasio, 2022) (Damasio and Damasio, 2023)	Affective origin of consciousness	Homeostasis	Feelings are foundational to conscious experience and adaptive behavior.
(Baars, 1993)	Global Workspace Theory (GWT)	Prefrontal-parietal networks	Consciousness emerges through information broadcasting ("ignition").
(Tononi et al., 2016)	Integrated Information Theory (IIT)	Network structure	Consciousness is defined by high Φ (integration + differentiation).
(Clark, 2016) (Seth, 2021)	Predictive Processing (PP)	Cortical hierarchy	Consciousness as controlled hallucination; inference minimizes prediction error.
(Kok et al., 2012)	fMRI of predictive signals	Visual cortex (V1)	Prefrontal predictions modulate V1 activity before stimulus onset.
(Bastos et al., 2015)	MEG and prediction error	Oscillatory activity	Prediction error associated with gamma desynchronization.
(Lamme, 2006)	Recurrent Processing Theory (RPT)	Feedback loops	Consciousness requires recurrent top-down processing.
(Rosenthal, 2008)	Higher-Order Thought (HOT) Theory	Metacognition	Awareness arises from mental representations of mental states.
(Dennett, 1991)	"Cartesian Theater" critique	Phenomenology	Critiques central observer; proposes distributed narrative processing.
(Rensink, 1997) (Blackmore, 2002)	Change/inattentional blindness	Visual perception	Conscious perception is sparse and filled in by expectation.
(Gazzaniga, 1992)	Interpreter of left hemisphere	Split-brain studies	Left hemisphere creates a narrative illusion of a unified self.
(Libet et al., 1983)	Readiness potential preceding decision	Volition studies	Brain activity precedes conscious decision to act.
(Wegner, 2003)	Illusion of conscious will	Agency construction	Agency is inferred post hoc from predictive coherence.
(Birch et al., 2020)	Animal consciousness framework	Comparative cognition	Proposes five axes for gradated consciousness (e.g., unity, temporality).
(Storm et al., 2024)	Integrative consciousness model	Multiscale theory	Hierarchical model from homeostasis to

			metacognition and global ignition.
(Aru et al., 2012)	Dendritic integration in conscious processing	Apical dendrites	Cortical feedback and sensory input integration support conscious perception.
(Hameroff and Penrose, 2014)	Orch-OR theory (quantum consciousness)	Microtubules	Speculative link between quantum coherence and awareness.
(Vitiello, 2020)	Quantum field theory of mind	Dissipative systems	Consciousness as a field-theoretical emergent phenomenon.
(Nagel, 1980)	"What is it like to be a bat?"	Phenomenology	Subjective experience cannot be reduced to objective scientific terms.

References

Alcaraz-Sánchez A. Minimal states of awareness across sleep and wakefulness: A multidimensional framework to guide scientific research. *Philosophy and the Mind Sciences*. 2024;5. doi:10.33735/phimisci.2024.10233

Aru J, Bachmann T, Singer W, Melloni L. Distilling the neural correlates of consciousness. *Neurosci Biobehav Rev*. 2012;36(2):737-746. doi:10.1016/j.neubiorev.2011.12.003

Baars BJ. *A Cognitive Theory of Consciousness*. Cambridge University Press; 1993.

Bartlett G. Does integrated information theory make testable predictions about the role of silent neurons in consciousness? *Neurosci Conscious*. 2022;2022(1):niac015. doi:10.1093/nc/niac015

Bast A, Guerrero Olloqui D, Montijn N, et al. Thalamus enables active dendritic coupling of inputs arriving at different cortical layers. *Nat Commun*. 2025;16:8327.

Bastos AM, Litvak V, Moran R, et al. A DCM study of spectral asymmetries in feedforward and feedback connections between visual areas V1 and V4 in the monkey. *Neuroimage*. 2015;108:460-475. doi:10.1016/j.neuroimage.2014.12.081

Birch J, Schnell AK, Clayton NS. Dimensions of Animal Consciousness. *Trends Cogn Sci*. 2020;24(10):789-801. doi:10.1016/j.tics.2020.07.007

Blackmore S. There is no stream of consciousness. *Journal of Consciousness Studies*. 2002;9(5-6):17-28.

Block N. Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav Brain Sci*. 2007;30(5-6):481-499. doi:10.1017/S0140525X07002786

Camargo DK. Cartesian Dualism: Bridging Descartes' Mind-Body Dichotomy to Contemporary Psychology and Mental Health. *Journal of Contemporary Philosophical and Anthropological Studies*. 2025;3(2). doi:10.59652/jcpas.v3i2.513

Cavanna AE, Nani A. Colin McGinn: The Mysterious Flame of Consciousness. In: *Consciousness: Theories in Neuroscience and Philosophy of Mind*. Springer; 2014:55-59. doi:10.1007/978-3-662-44088-9_10

Chalmers DJ. Facing up to the problem of consciousness. *Journal of Consciousness Studies*. 1995;2(3):200-219.

Clark A. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press; 2016.

Damasio A, Damasio H. Homeostatic feelings and the biology of consciousness. *Brain*. 2022;145(7):2231-2235. doi:10.1093/brain/awac194

Damasio A, Damasio H. Feelings Are the Source of Consciousness. *Neural Comput*. 2023;35(3):277-286. doi:10.1162/neco_a_01521

Dehaene S, Naccache L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*. 2001;79(1-2):1-37. doi:10.1016/s0010-0277(00)00123-2

Dennett DC. *Consciousness Explained*. Little Brown; 1991.

Fischer DB, Boes AD, Demertzi A, et al. A human brain network derived from coma-causing brainstem lesions. *Neurology*. 2016;87(23):2427-2434. doi:10.1212/WNL.0000000000003404

Fuller PM, Sherman D, Pedersen NP, Saper CB, Lu J. Reassessment of the structural basis of the ascending arousal system. *J Comp Neurol*. 2011;519(5):933-956. doi:10.1002/cne.22559

Gazzaniga MS. *Nature's Mind: The Biological Roots of Thinking, Emotions, Sexuality, Language, and Intelligence*. Basic Books; 1992.

Hameroff S, Penrose R. Consciousness in the universe: A review of the 'Orch OR' theory. *Physics of Life Reviews*. 2014;11(1):39-78. doi:10.1016/j.plrev.2013.08.002

Honjoh S, Sasai S, Schiereck SS, Nagai H, Tononi G, Cirelli C. Regulation of cortical activity and arousal by the matrix cells of the ventromedial thalamic nucleus. *Nat Commun*. 2018;9(1):2100. doi:10.1038/s41467-018-04497-x

Hume D. *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method into Moral Subjects*. Library of Alexandria; 2020.

Irala D, Sakers K, Sabeur K, et al. Astrocyte-secreted neurocan controls inhibitory synapse formation and function. *Neuron*. 2024;112(10):1657-1675.e10. doi:10.1016/j.neuron.2024.03.007

Jang H, Mashour GA, Hudetz AG, Lee U. Measuring the dynamic balance of integration and segregation underlying consciousness, anesthesia, and sleep in humans. *Nat Commun*. 2024;15(1):9164. doi:10.1038/s41467-024-53299-x

Kerkhofs M, Lavie P. Frédéric Bremer 1892–1982: a pioneer in sleep research. *Sleep Med Rev*. 2000;4(5):505-510. doi:10.1053/smrv.2000.0112

Kok P, Rahnev D, Jehee JF, Lau HC, de Lange FP. Attention reverses the effect of prediction in silencing sensory signals. *Cereb Cortex*. 2012;22(9):2197-2206. doi:10.1093/cercor/bhr310

Lamme VAF. Towards a true neural stance on consciousness. *Trends Cogn Sci*. 2006;10(11):494-501. doi:10.1016/j.tics.2006.09.001

Levine J. Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*. 1983;64(4):354-361.

Libet B, Gleason CA, Wright EW, Pearl DK. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*. 1983;106(Pt 3):623-642. doi:10.1093/brain/106.3.623

Mckilliam A. A mechanistic alternative to minimal sufficiency as the guiding principle for NCC research. *Neurosci Conscious*. 2024;2024(1):niae014. doi:10.1093/nc/niae014

Nagel T. What is it like to be a bat? In: *The Language and Thought Series*. Harvard University Press; 1980:159-168.

Nichols TA. Anti-NMDA receptor encephalitis: An emerging differential diagnosis in the psychiatric community. *Ment Health Clin*. 2016;6(6):297-303. doi:10.9740/mhc.2016.11.297

Palencia JLD. Neuroentropy and Consciousness: A Meta-Analytic Review. *BRAIN: Broad Research in Artificial Intelligence and Neuroscience*. 2025;16(4). doi:10.70594/brain/16.4/16

Pinter C. *Mind and the Cosmic Order: How the Mind Creates the Features & Structure of All Things, and Why this Insight Transforms Physics*. Springer; 2020. doi:10.5325/GESTALTREVIEW.25.1.0143

Rensink RA, O'Regan JK, Clark JJ. To see or not to see: The need for attention to perceive changes in scenes. *Psychol Sci*. 1997;8(5):368-373. doi:10.1111/j.1467-9280.1997.tb00427.x

Rosenthal DM. Consciousness and its function. *Neuropsychologia*. 2008;46(3):829-840. doi:10.1016/j.neuropsychologia.2007.11.012

Schiffer F. The physical nature of subjective experience and its interaction with the brain. *Med Hypotheses*. 2019;125:57-69. doi:10.1016/j.mehy.2019.02.011

Seth A. *Being You: A New Science of Consciousness*. Faber & Faber; 2021.

Shea N, Frith CD. The Global Workspace Needs Metacognition. *Trends Cogn Sci*. 2019;23(7):560-571. doi:10.1016/j.tics.2019.04.007

Son Y. Molecular mechanisms of general anesthesia. *Korean J Anesthesiol.* 2010;59(1):3-8. doi:10.4097/kjae.2010.59.1.3

Storm JF, Klink PC, Aru J, Senn W, Goebel R, Singer W. An integrative, multiscale view on neural theories of consciousness. *Neuron.* 2024;112(10):1531-1552. doi:10.1016/j.neuron.2024.02.004

Sumi T, Harada K. Mechanism underlying hippocampal long-term potentiation and depression based on competition between endocytosis and exocytosis of AMPA receptors. *Sci Rep.* 2020;10(1):14711. doi:10.1038/s41598-020-71528-3

Teleanu RI, Niculescu AG, Roza E, Vladâcenco O, Grumezescu AM, Teleanu DM. Neurotransmitters-Key Factors in Neurological and Neurodegenerative Disorders of the Central Nervous System. *Int J Mol Sci.* 2022;23(11):5954. doi:10.3390/ijms23115954

Tononi G, Boly M, Massimini M, Koch C. Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci.* 2016;17(7):450-461. doi:10.1038/nrn.2016.44

Vitiello G. Matter, mind and consciousness: from information to meaning. *J Integr Neurosci.* 2020;19(4):701-709. doi:10.31083/j.jin.2020.04.310

Wegner DM. *The Illusion of Conscious Will.* MIT Press; 2003.

Weissman DG, Nook EC, Dews AA, et al. Low Emotional Awareness as a Transdiagnostic Mechanism Underlying Psychopathology in Adolescence. *Clin Psychol Sci.* 2020;8(6):971-988. doi:10.1177/2167702620923649

Yon D. *A Trick of the Mind: How the Brain Invents Your Reality*. CornerStone Press; 2025.

Zhuo M. Cortical synaptic basis of consciousness. *Eur J Neurosci.* 2024;59(5):796-806. doi:10.1111/ejn.16198