Network Biology, 2026, 16(4): 552-585

Article

The Feedback Integration Theory of Consciousness: A unified neurobiological framework

WenJun Zhang
School of Life Sciences, Sun Yat-sen University, Guangzhou, China
E-mail: zhwj@mail.sysu.edu.cn, wjzhang@iaees.org

Received 3 May 2026; *Accepted 10 May 2026*; *Published online 15 May 2026*; *Published 1 December 2026*

Abstract

Consciousness remains one of the most profound and challenging phenomena confronting modern science. Despite decades of empirical progress across multiple levels of analysis—from molecular pharmacology to systems neuroscience to comparative cognition—no single theoretical framework has successfully unified the diverse and often contradictory lines of evidence into a coherent, mechanistically explicit account of subjective experience. This paper undertakes a comprehensive, multi-level review of the biological foundations of consciousness, systematically examining a wide range of empirical and theoretical literature across seven interconnected domains: the definitional and operational foundations of consciousness research, the major competing theoretical frameworks and their evidential bases, the neuroanatomical structures and network dynamics causally implicated in conscious states, the comparative and evolutionary distribution of consciousness-related capacities across phylogenetically distant taxa, the causal-manipulative evidence from anesthesia and neurostimulation studies, the identification of persistent anomalies that challenge existing theories, and the systematic construction of a novel theoretical framework that resolves these anomalies. Building upon this extensive synthesis, I propose the Feedback Integration Theory (FIT), a novel theoretical framework grounded in the convergent empirical finding that the neural signature most consistently associated with conscious experience across paradigms, modalities, and species is the recurrent, feedback-driven integration of information within hierarchically organized neural architectures. The theory makes four central claims, each supported by converging lines of evidence from multiple experimental traditions: (1) The minimal physical substrate of a conscious percept—the neural correlate of phenomenal consciousness—is a recurrently connected cortical (or cortical-homologous) loop capable of sustaining temporally structured feedback processing. The feedforward sweep of activation, regardless of its complexity or anatomical extent, is insufficient to generate subjective experience; consciousness requires the re-entrant modulation of earlier processing stages by later ones. (2) The specific qualitative character (quale) of a conscious experience is determined by the particular spatiotemporal pattern of feedback convergence across representational layers—the "feedback signature" that is unique to each conscious content. (3) The distinction between phenomenal consciousness (the subjective "what-it-is-like") and access consciousness (the availability of content for report, reasoning, and behavioral control) maps onto the neuroscientific distinction between local recurrent processing within sensory-specific hierarchies and global broadcasting via fronto-parietal networks, respectively. (4) The unified, field-like character of conscious experience—the binding of diverse features into a single coherent scene—emerges from phase-coupled feedback coordination across multiple specialized processing loops, orchestrated by thalamocortical resonance mechanisms centered on the intralaminar thalamic nuclei. FIT provides principled, mechanistically explicit resolutions to several persistent anomalies that have resisted explanation by existing theories. These include the phenomenon of phenomenal consciousness without cognitive access, the cerebellum puzzle (why a structure with enormous computational capacity and ordered circuitry does not generate reportable conscious experience), the differential effects of pharmacologically distinct anesthetic agents that nonetheless converge on loss of consciousness, and the convergent evolution of consciousness-supporting neural architectures in phylogenetically distant taxa with radically different brain organizations. The theory generates a suite of precise, falsifiable, and empirically testable predictions amenable to current experimental techniques including optogenetics, high-density electrophysiology, and layer-specific neuroimaging. Finally, the paper addresses the "hard problem" of consciousness from a principled physicalist perspective, arguing that subjective experience is not an additional property requiring explanation beyond the physical, but rather constitutes the intrinsic, first-person perspective of a physical system's own causal integration structure—the way a sufficiently complex recurrent system "feels" from the inside.

Keywords consciousness; feedback integration; recurrent processing; neural correlates of consciousness; phenomenology; thalamocortical circuits; anesthesia; integrated information; global workspace; comparative; cognition; evolutionary neuroscience; hard problem of consciousness.

Network Biology
ISSN 2220-8879
URL: http://www.iaees.org/publications/journals/nb/online-version.asp
RSS: http://www.iaees.org/publications/journals/nb/rss.xml
E-mail: networkbiology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 The Phenomenon of Consciousness: Precise Definition and Operationalization

The scientific investigation of any natural phenomenon requires, as its foundational prerequisite, a precise, unambiguous, and operationally tractable definition of the object of study. Consciousness, long considered the exclusive province of philosophical inquiry, has in recent decades become a legitimate target of empirical neuroscience, but its scientific maturation depends crucially on the clarity with which the phenomenon itself is delineated. A rigorous analysis of the conceptual landscape reveals that "consciousness" is not a unitary, monolithic construct, but rather comprises multiple inter-related yet empirically dissociable dimensions that must be carefully distinguished to avoid conceptual conflation and experimental ambiguity (Zeman, 2008; Malcom, 2023; Zhang, 2027).

1.1 The Multidimensional Architecture of Consciousness

The fundamental distinction in the clinical and experimental neuroscience of consciousness is between the level of consciousness (wakefulness or arousal) and the contentof consciousness (awareness of specific stimuli, thoughts, or emotions). This bidimensional framework, articulated with increasing precision over the past two decades, provides a powerful taxonomy for differentiating a wide range of healthy and pathological states (Zeman, 2008). The two dimensions are positively correlated in normal physiological states—both are high during focused wakefulness and low during deep slow-wave sleep—yet they can dissociate dramatically under pathological and pharmacological conditions, providing unique windows into the neural substrates that support each dimension independently.

Arousal, or vigilance, refers to the global level of brain activation that determines an organism's capacity for any conscious processing whatsoever. This dimension is primarily sustained by the ascending reticular activating system (ARAS), a collection of brainstem nuclei that project diffusely to thalamic and cortical targets via both cholinergic and monoaminergic pathways. The brainstem reticular formation was first identified as critical for consciousness in the landmark experiments of Moruzzi and Magoun, who demonstrated in 1949 that electrical stimulation of the rostral brainstem reticular formation produced a generalized, non-specific activation of the cerebral cortex, while lesions to this region resulted in persistent coma. As Plets (1998) summarized in a comprehensive anatomical review, "consciousness is based on the anatomical and physiological integrity of the ascending reticular activating system and its widespread connections with the major somatic and special sensory pathways." Stimulation of the ARAS provoked diffuse EEG desynchronization and behavioral arousal, establishing the fundamental principle that the brainstem mechanisms regulating arousal are anatomically and functionally dissociable from the cortical mechanisms that support the specific contents of consciousness (Plets, 1998).

Studies of disorders of consciousness have further refined our understanding of the arousal system. Arousal is now known to be regulated by lower-order regions of the brain—the brainstem's ascending reticular activating system, thalamus, hypothalamus, and basal forebrain—collectively designated the subcortical arousal network (Yang et al., 2025). This network exerts its effects through multiple neurotransmitter systems: serotonin secreted by the dorsal raphe nucleus activates default mode network nodes; norepinephrine secreted by the locus coeruleus regulates mood, attention, and motivation; dopamine secreted by the ventral tegmental area serves as a connectivity hub between subcortical arousal and cortical awareness networks; and the laterodorsal and pedunculopontine tegmental nuclei regulate the sleep-wake cycle (Yang et al., 2025). The Harvard Ascending Arousal Network Atlas has provided detailed structural characterization of these pathways, confirming that "AAN neurons connect the brainstem to the thalamus, hypothalamus, basal forebrain and cortex, activating cortically based awareness networks" (Brian and Hannah, 2023).

Phenomenal consciousness, the dimension of subjective experience itself—"what it is like" to perceive a red rose, to feel pain, or to experience an emotion—was systematically distinguished from other dimensions of consciousness by the philosopher Ned Block. Block (1995) designated this dimension as phenomenal consciousness (P-consciousness) to distinguish it from access consciousness (A-consciousness), the availability of mental content for the rational control of thought and action, including verbal report, deliberate reasoning, and the guidance of voluntary behavior. This conceptual distinction has proven to be of profound empirical significance, generating an extensive experimental literature examining whether P-consciousness can occur in the absence of cognitive access—a question that strikes at the heart of competing theories of consciousness (Block, 2007; Usher et al., 2018).

The neural correlates of consciousness (NCC) research program, as formalized by Crick and Koch (1990) and elaborated by subsequent investigators, seeks to identify "the minimal set of neuronal events sufficient for any one specific conscious percept" (Koch et al., 2016). This research program has generated substantial empirical progress, but its interpretation depends critically on which dimension of consciousness is being targeted. As Hohwy (2007) noted in a comprehensive methodological review, "the science of consciousness begins with the search for the neural correlates of consciousness," but "different conceptions of consciousness may influence the search for the neural correlates," including "the distinctions between access and phenomenal consciousness, and between conscious states and unified conscious fields." The REF (reorganization of elementary functions) framework proposed by Mogensen and Overgaard (2018) conceptualizes the relationship between phenomenal and access consciousness within a neurocognitive model in which "phenomenal consciousness cannot 'overflow' availability of information for action," arguing instead that "phenomenal consciousness may overflow working memory because working memory is seen as a surface phenomenon reflecting underlying dynamic strategies influenced by both experience and situational factors."

1.2 The Arousal–Awareness Taxonomy and Its Clinical Dissociations

The arousal–awareness bidimensional model provides a powerful framework for differentiating a wide range of healthy and pathological states. Zeman (2008) demonstrated that arousal and awareness are positively correlated in normal physiological states—both are high during focused wakefulness and low during deep slow-wave sleep—with the notable exception of rapid eye movement (REM) sleep, where awareness in the form of vivid dreaming can be high despite behavioral quiescence. This dissociation during REM sleep provides a crucial natural experiment: consciousness can be rich and immersive in the near-complete absence of sensory input or motor output, indicating that consciousness is fundamentally an intrinsically generated phenomenon of the activated brain rather than a simple reflection of sensory processing or behavioral engagement (Pang, 2023).

Pathological conditions reveal further dissociations that illuminate the anatomical and functional architecture of consciousness. Coma is a state of complete unarousability with absent sleep-wake cycles, resulting from either bilateral hemispheric damage or brainstem lesions affecting the ARAS. The vegetative state, now more neutrally termed unresponsive wakefulness syndrome, is characterized by preserved sleep-wake cycles and spontaneous eye opening in the complete absence of behavioral evidence of awareness—a dissociation that highlights the anatomical separation between structures supporting arousal and those supporting awareness. The minimally conscious state (MCS) represents an intermediate condition in which patients demonstrate inconsistent but clearly discernible behavioral evidence of awareness, such as visual pursuit, localization to noxious stimulation, or contingent vocalization. The locked-in syndrome, by contrast, involves preserved consciousness with near-complete paralysis, underscoring the independence of conscious experience from motor output (Giacino et al., 2014).

As described by a team of researchers led by Professor Yang from Capital Medical University in China, "arousal is regulated by lower-order regions of the brain—the brainstem's ascending reticular activating system (ARAS), thalamus, hypothalamus, and basal forebrain—collectively called the subcortical arousal network (SAN). Whereas awareness is controlled by cortical circuits such as the frontal-parietal network (FPN) that responds to external stimuli, and the default mode network (DMN) that is aware of internal stimuli. Together, this is known as the cortical awareness network (CAN). Interruptions to arousal networks can result in a coma, whereas disruptions to awareness circuitry can result in a vegetative state or minimally conscious state" (Yang et al., 2025).

1.3 Operationalizing Consciousness Across Species, States, and Modalities

A fundamental challenge for the science of consciousness is the development of reliable operational definitions that do not depend solely on verbal report. This is essential for extending consciousness research to preverbal infants, non-human animals, and patients with disorders of consciousness who may retain some awareness despite their inability to communicate. Several converging methodological approaches have been developed, each with distinct strengths and limitations.

The multidimensional framework for understanding interspecies variation in states of consciousness proposed by Birch et al. (2020) distinguishes five key dimensions of variation relevant to animal consciousness research: perceptual richness, evaluative richness, integration at a time, integration across time, and self-consciousness. This framework explicitly recognizes that "a key goal for animal consciousness research should be to produce a much richer evidence base for the construction of consciousness profiles and more precise ways of measuring the dimensions," moving beyond the simplistic question of whether a species "has" consciousness to a more nuanced, graded assessment of which dimensions of consciousness are present and to what degree (Birch et al., 2020).

Mirror self-recognition has been proposed as a behavioral marker of self-awareness across species.

Originally developed by Gallup (1970), the mirror test assesses whether an animal can recognize its own reflection as an image of itself rather than another conspecific. MSR has been demonstrated in all great ape species, dolphins, elephants, and more recently in cleaner fish, which can identify their mirror reflection via self-face recognition (Gallup, 1977). The evolutionary significance of this capacity is underscored by the self-awareness homology hypothesis, which proposes that MSR and its underlying cognitive architecture trace back to a common ancestor of vertebrates. However, the interpretation of MSR as a definitive marker of self-consciousness remains contested, and its absence in a given species does not necessarily indicate the absence of self-awareness.

The New York Declaration on Animal Consciousness, signed by 288 researchers and reviewed by numerous studies conducted over the past decade, suggests "that consciousness is now a 'realistic' possibility not only in birds and mammals but also in reptiles, amphibians, fish, cephalopods like octopuses and cuttlefish, crustaceans like hermit crabs and crayfish, and insects like bees and fruit flies." The declaration focuses on "phenomenal consciousness" or "sentience," defined as the capacity for subjective experiences, "for example, the experience of a particular touch, taste, sight, or smell" and "experiences that feel good or bad, for example, pleasure, pain, hope, and fear" (Özen et al., 2024).

The most rigorously validated neurophysiological metric of consciousness that is independent of behavioral report is the Perturbational Complexity Index (PCI), developed by Casali et al. (2013). This measure quantifies the spatiotemporal complexity of the brain's electrophysiological response to transcranial magnetic stimulation (TMS). As described by Casali et al. (2013), "PCI is calculated by (i) perturbing the cortex with transcranial magnetic stimulation (TMS) to engage distributed interactions in the brain (integration) and (ii) compressing the spatiotemporal pattern of these electrocortical responses to measure their algorithmic complexity (information)." In conscious states, TMS triggers a complex, rapidly changing pattern of cortical activation that propagates widely across the brain; in states of reduced consciousness, such as deep sleep, anesthesia, or vegetative state, the same stimulation evokes a simple, stereotyped, and locally confined response. Critically, "PCI reliably discriminated the level of consciousness in single individuals during wakefulness, sleep, and anesthesia, as well as in patients who had emerged from coma and recovered a minimal level of consciousness" (Casali et al., 2013). The PCI was lowest in patients in a vegetative state, falling within values of healthy people in states of deep sleep and anesthesia, while patients in minimally conscious state showed intermediate values, and locked-in patients showed values comparable to fully conscious controls (Casali et al., 2013).

The non-equilibrium dynamics of the unperturbed brain have also been shown to predict consciousness level. Stikvoort et al. (2025) demonstrated "that states of consciousness involving lower arousal and/or lower awareness had a lower level of asymmetry in their effective connectivities, a lower level of irreversibility in their simulated dynamics, and a lower complexity compared to control subjects," showing that "the asymmetry in the underlying connections drives the nonequilibrium state of the system and in turn the differences in complexity as a response to external stimuli."

The compatibility between PCI and the Global Neuronal Workspace Theory has been explored by Farisco and Changeux (2023), who argued that "the cortical response to TMS has been interpreted as an 'index' of the level of consciousness" that can be accommodated within the GNWT framework as a measure of the brain's capacity for broadcasting information to a large-scale network.

2 Major Theoretical Frameworks: Systematic Review and Critical Analysis

The contemporary neuroscience of consciousness is characterized by a plurality of competing theoretical frameworks, each offering distinct mechanistic accounts of how subjective experience arises from neural activity, and each supported by partially overlapping yet partially divergent bodies of empirical evidence. A comprehensive understanding of the biological foundations of consciousness requires a systematic examination of the major theories, their core claims, their evidential bases, and—crucially—the empirical anomalies that each fails to adequately explain. As described by Maillé and Lynn (2020), "Understanding the neural basis of consciousness is one of the fundamental challenges in modern neuroscience," and "a number of sophisticated models and theories have attempted to formalize how the brain implements consciousness using insights from philosophy, psychology, computer science, and neuroscience."

2.1 The Global Neuronal Workspace Theory

The Global Neuronal Workspace (GNW) theory, originally proposed as a psychological model by Baars (1988) and subsequently elaborated into a detailed neurobiological framework by Dehaene et al. (Dehaene & Changeux, 2011; Dehaene et al., 2003), posits that consciousness arises from the global broadcasting and sustained amplification—termed "ignition"—of information across a distributed network of interconnected cortical regions. According to this framework, "at any given time, many modular cerebral networks are active in parallel and process information in an unconscious manner. An information becomes conscious when it gains access to a global neuronal workspace, a distributed network of pyramidal neurons with long-distance connections, particularly dense in prefrontal, parietal, and cingulate regions" (Dehaene et al., 2006). The core computational operation posited by GNW is that conscious access involves "a distribution of sensory information from sensory cortices to parietal and prefrontal cortex" (Dehaene & Naccache, 2001; Baars, 2002), with increased activity in frontoparietal regions under conscious perception seen as evidence for such a 'broadcasting' of sensory information (Lara and Wallis, 2015).

The empirical evidence supporting GNW is substantial. Electrophysiological studies have consistently demonstrated that consciously perceived stimuli elicit a characteristic late positive component in event-related potentials—the P3b, peaking around 300-500 milliseconds after stimulus onset—which is absent or greatly reduced for stimuli that do not reach conscious awareness. Neuroimaging studies have shown that conscious perception, compared with unconscious processing of the same stimuli, is associated with increased activation across a distributed fronto-parietal network, as well as increased functional connectivity between these regions and sensory cortices (Dehaene & Changeux, 2011). The theory further postulates that "this global availability of information through the workspace is what we subjectively experience as a conscious state" (Dehaene et al., 2006).

However, GNW faces several significant empirical challenges. The role of the prefrontal cortex (PFC) in conscious experience has been questioned by both clinical and experimental evidence. As reviewed by Koch et al. (2016), "recent findings showing that the anatomical neural correlates of consciousness are primarily localized to a posterior cortical hot zone that includes sensory areas, rather than to a fronto-parietal network involved in task monitoring and reporting." Patients with extensive bilateral prefrontal lesions can retain seemingly intact conscious experience despite profound deficits in executive function and cognitive control. Moreover, in an fMRI study designed to distinguish neural correlates of conscious content from those of attention and report, researchers "found no evidence for changes in prefrontal representation of sensory information under increased levels of awareness," suggesting that increased activity in frontoparietal regions with increased awareness might simply reflect unspecific processes, say as in detecting or reporting a change in perception, rather than coding the sensory information itself (Lara and Wallis, 2015).

In a landmark adversarial collaboration directly juxtaposing IIT and GNWT with 256 human participants, the COGITATE consortium (Cogitate Consortium et al., 2025) found that "GNWT is challenged by the general lack of ignition at stimulus offset and limited representation of certain conscious dimensions in the prefrontal cortex." These findings substantially challenge key tenets of GNWT and suggest that the theory's emphasis on prefrontal ignition as the sine qua non of consciousness may require significant revision (Cogitate Consortium et al., 2025).

2.2 Integrated Information Theory

Integrated Information Theory (IIT), developed by Giulio Tononi and colleagues (Tononi, 2004, 2008; Tononi et al., 2016; Oizumi et al., 2014), takes a fundamentally different approach. Rather than identifying consciousness with a particular cognitive function such as global broadcasting, IIT defines consciousness in terms of the intrinsic causal properties of a physical system. The theory "starts from experience itself via five phenomenological axioms: intrinsic existence, composition, information, integration, and exclusion" (Tononi & Koch, 2015) and derives from these axioms a mathematical measure, Φ, which quantifies the amount of integrated information generated by a system—the degree to which the system's causal interactions specify a cause-effect structure that is irreducible to the interactions of its parts. According to IIT, "consciousness is a fundamental property possessed by physical systems having specific causal properties. It predicts that consciousness is graded, is common among biological organisms, and can occur in some very simple systems" (Tononi & Koch, 2015).

The neuroanatomical predictions of IIT have identified what its proponents term the "posterior hot zone"—a hub involving parietal, occipital, and temporal cortical regions—as playing "a key role in controlling information integration by the full NCC" (Boly et al., 2017; Tononi et al., 2016). According to this framework, "the anatomical neural correlates of consciousness are primarily localized to a posterior cortical hot zone that includes sensory areas, rather than to a fronto-parietal network involved in task monitoring and reporting" (Koch et al., 2016). The posterior hot zone, particularly regions such as the precuneus, posterior cingulate, and retrosplenial cortices, has been identified as "pivotal for conscious information processing," with evidence drawn from "posterior cingulate epilepsy cases, midcingulate lesions that de-afferent this region and are associated with unilateral sensory neglect, observations from stroke and vegetative state patients, alterations in blood flow during sleep, and the actions of anesthetics" (Vogt and Laureys, 2005).

IIT makes specific predictions about which neural architectures can and cannot support consciousness. According to the theory, "the physical substrate of consciousness—constituted of a large number of neurons arranged as 'pyramids of grids' and located predominantly in posterior cortical areas—can specify a cause-effect structure with high Φ" during wakefulness (Tononi et al., 2016). By contrast, "when eight units are connected pairwise as parallel modules with minimal intermodular connectivity, the network specifies four separate, minimal cause-effect structures, each with very low Φ. Modular connectivities are found, for instance, in the cerebellum". Thus, "the paradoxes posed by cerebellum, the basal ganglia, and by other large neuronal aggregates that seem irrelevant for consciousness, are less disconcerting; their architecture is not suited for integrating information, whereas the thalamocortical system seems to contain a core of high Φ" (Tononi, 2012).

However, IIT faces several major challenges. The calculation of Φ for any realistic neural system is computationally intractable, making direct empirical tests of the theory's quantitative predictions extremely difficult. Critics have also argued that "some proponents of the Integrated Information Theory (IIT) of consciousness profess strong views on the Neural Correlates of Consciousness (NCC), namely that large swathes of the neocortex, the cerebellum, at least some sensory cortices, and the so-called limbic system are all not essential for any form of conscious experiences. I argue that this connection is not incidental. Conflation between strong and weak versions of the theory has led these researchers to adopt definitions of NCC that are inconsistent with their own previous definitions" (Michel & Lau, 2020).

The most telling challenge to IIT comes from the adversarial collaboration mentioned above. The COGITATE consortium (Cogitate Consortium et al., 2025) found that "for IIT, a lack of sustained synchronization within the posterior cortex contradicts the claim that network connectivity specifies consciousness." This finding directly challenges a core prediction of IIT regarding the nature of the neural substrate of consciousness (Cogitate Consortium et al., 2025).

2.3 Higher-Order Theories

Higher-order theories (HOT) of consciousness constitute a distinct family of theoretical approaches that share the central claim that a mental state becomes conscious only when it is represented by a higher-order mental state—that is, when the brain forms a representation about that state. As articulated by Lau and Rosenthal (2011), "higher-order theories of consciousness argue that conscious awareness crucially depends on higher-order mental representations that represent oneself as being in particular mental states." In the higher-order thought version of the theory, this higher-order representation takes the form of a conceptual, belief-like mental state; in the higher-order perception variant, it takes the form of a quasi-perceptual meta-representation.

The neural instantiation of higher-order theories posits that "a sensory representation remains unconscious until it is mapped by a higher-order representation generated in prefrontal and metacognitive regions" (Brown et al., 2019). This framework accounts for the introspective dimension of consciousness—the fact that conscious experiences seem to be experiences that we are aware of having—and provides a natural explanation for why conscious contents are typically available for verbal report and cognitive control. As reviewed by Brown et al. (2019), "we propose a reconceptualization of lower-order states that contribute to higher-order awareness, including states of prefrontal cortex and multimodal and mnemonic states processed in posterior cortical areas."

The empirical evidence supporting higher-order theories includes findings that prefrontal activity correlates with conscious perception even when report-related and attentional confounds are controlled for, that metacognitive sensitivity (the ability to judge the accuracy of one's own perceptual decisions) is associated with prefrontal gray matter volume, and that disruption of prefrontal function through TMS can affect conscious perception. However, HOT faces challenges from evidence that extensive prefrontal lesions do not necessarily abolish conscious experience, and that rich phenomenal consciousness may be possible even in the absence of higher-order metacognitive access (Block, 2007).

2.4 Recurrent Processing Theory

The Recurrent Processing Theory, most extensively developed by Victor Lamme et al. (Lamme, 2000, 2003, 2006; Lamme & Roelfsema, 2000), occupies a unique position in the theoretical landscape. This theory directly addresses the neural mechanisms that distinguish conscious from unconscious visual processing, proposing that "the feed-forward activation of cells in any area of the brain is not sufficient to generate awareness, but that recurrent processing, mediated by horizontal and feedback connections is necessary" (Lamme, 2000). The theory distinguishes between: (1) the initial feedforward sweep of activation that propagates rapidly from primary sensory areas through the cortical hierarchy, which can support complex unconscious processing; (2) local recurrent processing within sensory cortices, mediated by horizontal and feedback connections, which generates phenomenal awareness; and (3) widespread recurrent processing involving fronto-parietal networks, which enables access consciousness—the ability to report, reason about, and act upon conscious content.

As Lamme (2003) argued, "phenomenal 'seeing' arises from local recurrent interactions in visual areas, while access-level 'knowing' depends on widespread recurrent activity involving fronto-parietal networks." This distinction between phenomenal and access consciousness within a unified recurrent processing framework provides a neuroscientifically grounded resolution to the debate about whether consciousness overflows cognitive access: phenomenal consciousness can occur whenever local recurrent processing is established, regardless of whether the content is subsequently globally broadcast and made available for report.

A comprehensive review by Tapia and Beck (2014) examined evidence for "probing feedforward and feedback contributions to awareness with visual masking and transcranial magnetic stimulation," noting that "a number of influential theories posit that visual awareness relies not only on the initial, stimulus-driven (i.e., feedforward) sweep of activation but also on recurrent feedback activity within and between brain regions." Recordings in subjects performing a substitution masking paradigm provided "direct support for the notion that recurrent processing in V1 correlates with visual awareness and that attention and awareness involve distinct recurrent processing operations" (Fahrenfort et al., 2008).

The primary evidence supporting recurrent processing theory comes from visual masking paradigms, in which a briefly presented target stimulus can be rendered invisible by a subsequently presented mask. Electrophysiological recordings in both humans and non-human primates have demonstrated that the feedforward activation of visual cortex proceeds normally for masked (unseen) stimuli, but the subsequent recurrent feedback activity is disrupted or abolished (Lamme et al., 2002). Similarly, in binocular rivalry paradigms, neurons throughout the visual hierarchy respond to both the consciously perceived and the suppressed stimulus during the initial feedforward sweep, but sustained recurrent activity correlates specifically with the conscious percept (Tong et al., 1998).

The recurrent processing theory also provides elegant explanations for the phenomenon of blindsight, in which patients with primary visual cortex (V1) lesions can perform above-chance visual discriminations despite reporting no conscious visual experience. The absence of V1 prevents the establishment of recurrent processing loops between extrastriate cortex and V1 that are necessary for phenomenal awareness, even though feedforward pathways through the superior colliculus and pulvinar can support unconscious visual processing (Lamme, 2000; Silvanto et al., 2005).

2.5 Evidence-Contradiction Matrix

A systematic comparison of the major theories reveals a landscape of partial convergence and persistent divergence. The following matrix summarizes the explanatory scope and empirical challenges of each major theoretical framework, identifying phenomena that constitute critical tests for distinguishing among theories.

Facts explained by most or all theories:

- Conscious perception is associated with more widespread and sustained neural activity than unconscious processing.
- The thalamocortical system is the primary neural substrate of consciousness; subcortical structures alone are insufficient.
- Information integration and differentiation are both features of conscious neural states.
- Attentional selection strongly influences which contents become conscious.
- Conscious contents are available for flexible behavioral control.

Facts that differentiate theories:

- Consciousness without report:Whether phenomenal consciousness can occur in the absence of fronto-parietal ignition and cognitive access (supports recurrent processing theory, challenges GNW and some versions of HOT).
- Prefrontal lesions and consciousness:Whether extensive bilateral prefrontal damage preserves consciousness (challenges GNW and HOT, consistent with recurrent processing and IIT).
- Cerebellum and consciousness:Why the cerebellum's enormous computational capacity does not generate reportable consciousness (explained by IIT through modular architecture, and by recurrent processing through lack of sustained recurrent loops; less naturally explained by GNW, which would need to specify why cerebellar processing is not globally broadcast).

- Posterior vs. anterior dominance:Whether the primary NCC is in posterior sensory regions (IIT, recurrent processing) or fronto-parietal networks (GNW, some HOT variants).
- Anesthetic convergence:Why pharmacologically diverse anesthetics all lead to loss of consciousness (explained by recurrent processing through disruption of feedback loops, by IIT through reduction of Φ, and by GNW through prevention of ignition).

Anomalous phenomena not adequately explained by any current theory:

- The specific qualitative character (quale) of different conscious experiences.
- The binding of diverse features into a unified conscious scene.
- The apparent unity of consciousness despite distributed neural processing.
- The transition from unconscious to conscious processing at a specific temporal threshold.
- The evolutionary origins and phylogenetic distribution of consciousness.

3 Neuroanatomical Necessities and Dynamic Sufficiencies

The identification of brain structures that are causally necessary for consciousness—structures whose damage or inactivation invariably leads to loss of consciousness—provides the strongest possible constraints on theories of consciousness. By analyzing the convergent evidence from lesion studies, disorders of consciousness, anesthesia research, and direct brain stimulation, we can distinguish between structures that are truly necessary for consciousness and those that are merely correlated with or modulatory of conscious states.

3.1 Anatomical Nodes of the Consciousness Network

The evidence from clinical neurology, neuroanatomy, and functional neuroimaging converges on a set of structures whose integrity is required for the maintenance of conscious states. These structures form a hierarchical network (Zhang, 2016a) in which lower-level structures enable the global state conditions for consciousness, while higher-level structures support the specific contents of conscious experience.

The Brainstem Ascending Reticular Activating System constitutes the foundational level of the consciousness hierarchy. As established by Moruzzi and Magoun (1949) and confirmed by decades of subsequent research, "damage to the upper brain stem reticular formation is known to cause the most radical disturbance of consciousness, i.e., coma" (Horn et al., 2020). The reticular formation is "a polysynaptic network, located centrally in the brainstem, that extends rostrally into the thalamus and hypothalamus and caudally into continuity with the propriospinal network of the spinal cord. The dendrites of this network are penetrated by the long ascending and descending specific fibre systems. As a consequence the ARAS is stimulated by the collaterals of every major somatic and special sensory pathway" (Plets, 1998). This anatomical organization explains why the ARAS is a necessary condition for consciousness and why its destruction results in irreversible coma, regardless of the integrity of cortical structures.

The neuromodulatory systems of the brainstem provide the chemical infrastructure for arousal. These include the serotonergic raphe nuclei, the noradrenergic locus coeruleus, the dopaminergic ventral tegmental area, the cholinergic pedunculopontine and laterodorsal tegmental nuclei, and the histaminergic tuberomammillary nucleus (Parvizi & Damasio, 2001). The diffuse projections of these systems to the thalamus and cortex regulate the excitability and signal-to-noise ratio of cortical neurons, creating the physiological conditions under which conscious processing can occur. The Harvard Ascending Arousal Network Atlas has mapped these pathways in detail, demonstrating that "AAN neurons connect the brainstem to the thalamus, hypothalamus, basal forebrain and cortex, activating cortically based awareness networks" (Brian and Hannah, 2023).

The Thalamus, particularly the intralaminar nuclei, occupies a critical position in the consciousness network. A systematic review by Cacciatore et al. (2025) examining the relationship between specific thalamic nuclei and consciousness found that "most of the retrieved positive evidence pertained to the intralaminar nuclear group, followed by the mediodorsal and ventral nuclear groups. Furthermore, when considering the specific nuclei within the intralaminar nuclear group, results highlighted the centromedian-parafascicular complex (CM-Pf) as the nucleus most related to consciousness" (Cacciatore et al., 2025). The review concluded that "though there is more research on the role of intralaminar nuclei, there is proportionally more positive evidence supporting these nuclei (particularly the CM-Pf) as key nodes in the network underlying consciousness compared to other thalamic nuclei." These findings "support ongoing therapeutic approaches to disorders of consciousness by reinforcing the rationale behind brain stimulation targeting CM-Pf" (Cacciatore et al., 2025).

The functional role of the higher-order thalamic nuclei in conscious perception has been directly demonstrated using simultaneous intracranial recordings from human thalamus and prefrontal cortex. Intralaminar and medial thalamic nuclei gate conscious perception and that transient thalamofrontal neural synchrony and cross-frequency coupling were both driven by the θ phase (2 to 8 Hz) of the activity in intralaminar and medial nuclei during the emergence of conscious perception. The coupling between these two thalamic nuclei and lateral PFC was stronger than that between these nuclei and other PFC subregions during conscious perception (Cacciatore et al., 2025; Fang et al., 2025). These findings provide direct sEEG evidence in the human brain that supports the gate role of the intralaminar and medial nuclei in the rapid process of conscious perception (Cacciatore et al., 2025; Fang et al., 2025).

The classical view of the thalamus as merely a relay station for sensory information has been substantially revised. The evidence now indicates that "the thalamus has been conventionally assumed to act as a prerequisite of sensory information for conscious perception rather than directly contributing to it. However, this assumption has been challenged by novel theoretical hypothesis and empirical evidence, mostly from anatomy and neuroimaging studies, which have shown the direct involvement of the human thalamus in various cognitive functions" (Cacciatore et al., 2025; Fang et al., 2025). Crick's hypothesis that "awareness is a function of reverberating corticothalamic loops and that the spotlight of intramodal attention is controlled by the reticular nucleus of the thalamus" (Smythies, 1997) has received substantial empirical support from these recent investigations.

The Posterior Cortical Hot Zonerepresents the cortical apex of the consciousness network. The comprehensive review by Koch et al. (2016) concluded that "the anatomical neural correlates of consciousness are primarily localized to a posterior cortical hot zone that includes sensory areas, rather than to a fronto-parietal network involved in task monitoring and reporting." This hot zone includes the precuneus, posterior cingulate cortex, retrosplenial cortex, and the temporo-parieto-occipital junction—regions that have been repeatedly identified as showing the most pronounced metabolic decreases during anesthesia, sleep, and disorders of consciousness.

The pivotal role of the posterior cingulate, precuneal, and retrosplenial cortices in consciousness has been "demonstrated with posterior cingulate epilepsy cases, midcingulate lesions that de-afferent this region and are associated with unilateral sensory neglect, observations from stroke and vegetative state patients, alterations in blood flow during sleep, and the actions of anesthetics" (Vogt and Laureys, 2005). These regions are uniquely positioned at the convergence of multiple functional networks, including the default mode network, the dorsal attention network, and the fronto-parietal control network, enabling them to serve as hubs for information integration across large-scale brain systems.

The posterior hot zone's role in consciousness is further supported by findings that "a large distributed cortical network would be necessary for conscious processes, involving prefrontal, cingulate and parietal regions," but that "the role of this anatomical network in consciousness remains a matter of debate: the prefrontal cortex could be excluded from the Neuronal Correlates of Consciousness (NCC), whereas a 'posterior hot zone' seems to play a predominant role in consciousness" (Martinaud, 2018).

3.2 Dynamic Signatures of the Conscious Brain

Beyond the anatomical structures that constitute the neural substrate of consciousness, the conscious brain is characterized by distinctive dynamic properties that differentiate it from unconscious states. These dynamic signatures provide not only diagnostic markers for the presence of consciousness but also mechanistic constraints on how neural activity gives rise to subjective experience.

The most robust and widely replicated dynamic signature of consciousness is the complexity of neural responses to perturbation. As described above, the PCI developed by Casali et al. quantifies the spatiotemporal complexity of TMS-evoked EEG responses and reliably discriminates conscious from unconscious states across a wide range of conditions. The theoretical basis for this measure is that "consciousness is defined as the combined presence of integration and segregation" in brain dynamics, and that "complexity emerges from the underlying critical dynamics" of the thalamocortical system (Stikvoort et al., 2025). The conscious brain operates at a critical point between order and chaos—a dynamical regime that maximizes both the differentiation of neural responses (information) and their integration into coherent patterns (Tononi et al., 2016).

A complementary dynamic signature is the level of non-equilibrium brain dynamics. Stikvoort et al. (2025) demonstrated "that states of consciousness involving lower arousal and/or lower awareness had a lower level of asymmetry in their effective connectivities, a lower level of irreversibility in their simulated dynamics, and a lower complexity compared to control subjects." The asymmetry in the underlying connections drives the non-equilibrium state of the system, which in turn generates the differences in complexity observed in response to external stimuli. This finding suggests that the conscious brain is fundamentally a non-equilibrium system, continuously driven away from thermodynamic equilibrium by the energy-consuming processes of neural activity, and that this non-equilibrium character is essential for the generation of complex, highly differentiated, yet integrated neural dynamics.

Functional connectivity and network dynamics (Zhang, 2010, 2012, 2016a-e, 2017, 2018) also differentiate conscious from unconscious states. During wakefulness, the brain exhibits high-dimensional, time-varying functional connectivity patterns; during deep sleep, anesthesia, or disorders of consciousness, functional connectivity becomes more stereotyped and low-dimensional. Conscious states are characterized by high temporal variability in functional connectivity, reflecting the brain's capacity to rapidly reconfigure its functional networks in response to changing internal and external demands. The transition to unconsciousness is associated with a collapse of this dynamic repertoire into a small number of dominant, repetitive patterns (Hudetz & Mashour, 2016).

The criticality hypothesis proposes that the brain operates near a phase transition between ordered and disordered dynamics, and that this critical regime optimizes information processing capacity. Tagliazucchi (2017) demonstrated that "the amount of integrated information (Φ) is maximal at the critical threshold," supporting "a link between information integration and conscious access." In this framework, "conscious awareness refers to information processing in the brain that is accompanied by subjective, reportable experiences" and the model "identifies conscious access with self-sustained percolation in an anatomical network" (Tagliazucchi, 2017).

3.3 The "Necessary-Sufficient" Layered Model

Synthesizing the anatomical and dynamic evidence, I propose a layered model of the neural prerequisites for consciousness that distinguishes between enabling conditions (necessary but not sufficient), core substrates (necessary and sufficient in combination for specific conscious contents), and modulatory influences (neither necessary nor sufficient individually but contributing to the richness and character of experience).

Layer 1: Enabling Conditions (Level of Consciousness)

- The brainstem ARAS, including its multiple neuromodulatory subsystems, provides the global activation state that makes conscious processing possible. Without ARAS function, no conscious content can be generated.
- The intralaminar thalamic nuclei, particularly the CM-Pf complex, provide the thalamocortical resonance that synchronizes cortical activity and enables the integration of information across distributed networks.

Layer 2: Core Substrates (Content of Consciousness)

- The posterior cortical hot zone, including the precuneus, posterior cingulate, and temporo-parieto-occipital junction, supports the specific contents of conscious experience through its capacity for high-dimensional information integration.
- Local recurrent processing loops within and between sensory-specific cortical regions generate the phenomenal character of specific conscious contents.
- Thalamocortical feedback loops, particularly involving higher-order thalamic nuclei, coordinate the binding of distributed features into unified conscious percepts.

Layer 3: Modulation and Access (Richness and Reportability)

- The fronto-parietal network, including dorsolateral prefrontal cortex and posterior parietal cortex, enables cognitive access to conscious contents—the ability to report, reason about, and act upon conscious experiences.
- Attentional systems, including the dorsal and ventral attention networks, select which of the many possible conscious contents gain access to the global workspace.
- Neuromodulatory systems fine-tune the balance between stability and flexibility in cortical dynamics, influencing the richness and character of conscious experience.

This layered model provides a framework for understanding the neural basis of consciousness that accommodates the dissociations observed in clinical conditions. The vegetative state represents a failure at Layer 1 (enabling conditions) with preserved structural integrity of Layers 2 and 3 but insufficient activation to support conscious processing. The minimally conscious state reflects partial but unstable activation of Layer 2. Frontal lobe lesions may impair Layer 3 (access and report) while leaving Layer 2 (phenomenal experience) largely intact—a prediction that aligns with the clinical observation that prefrontal patients can report conscious experiences despite profound executive deficits.

4 Comparative Cognition and the Evolutionary Origins of Consciousness

A comprehensive theory of the biological foundations of consciousness must be applicable not only to humans but to the diverse range of species in which consciousness-related capacities have been empirically documented. The comparative approach provides a crucial test of theoretical claims about the necessary and sufficient neural conditions for consciousness: if consciousness can arise in species with radically different neural architectures from those of mammals, then theories that identify consciousness with specific mammalian or primate brain structures must be revised.

4.1 The Phylogenetic Distribution of Consciousness Indicators

The evidence for consciousness-related capacities in non-human animals has expanded dramatically in recent decades, encompassing species from diverse taxonomic groups with markedly different neural organizations. This evidence base draws on multiple converging behavioral and neurobiological indicators, since no single behavioral test can definitively establish the presence of consciousness in a non-verbal organism.

Corvids and Parrots provide perhaps the most striking example of complex cognition and potential consciousness in a non-mammalian lineage. Pioneering research on avian cognition has highlighted that "avian species, mainly corvids and parrots, have a cognitive tool kit comparable with apes and other large-brained mammals, despite conspicuous differences in their neuroarchitecture" (Baciadonna and Cornero, 2021). New Caledonian crows manufacture and use tools with a sophistication rivaling that of chimpanzees; they can solve multi-step puzzles requiring causal reasoning about physical relationships; and they exhibit behaviors suggestive of future planning and episodic-like memory. The neural substrate for these capacities in birds is the pallium—particularly the nidopallium caudolaterale, which is considered functionally analogous to the mammalian prefrontal cortex despite having a radically different anatomical organization. The avian pallium is organized into nuclear clusters rather than the layered structure characteristic of mammalian neocortex, yet it supports comparably complex cognitive operations, demonstrating that "neural architectures distinct from primates' enable advanced cognitive abilities" (IBR, 2024).

Cetaceans present another compelling case of complex cognition in a lineage evolutionarily distant from primates. Bottlenose dolphins demonstrate mirror self-recognition, as shown by mirror-guided self-inspection behaviors, and exhibit complex social cognition including alliances, cooperative hunting, and cultural transmission of behaviors. The cetacean brain is characterized by a highly convoluted neocortex with a distinctive laminar organization, high numbers of Von Economo neurons (spindle cells previously thought to be unique to great apes and humans), and an elaborated paralimbic system. These features have evolved independently in cetaceans, representing a convergent solution to the computational demands of complex social cognition and conscious processing.

Cephalopods, particularly octopuses and cuttlefish, represent perhaps the most evolutionarily distant lineage for which compelling evidence of complex cognition exists. The cephalopod nervous system is organized on fundamentally different principles from that of vertebrates: it is distributed rather than centralized, with the majority of neurons located in the arms rather than the central brain, and it lacks the myelin sheaths that enable rapid signal propagation in vertebrates. Despite these radical differences, octopuses exhibit remarkable problem-solving abilities, observational learning, play behavior, and individual personality differences. The distributed neurological networks of cephalopods "enable remarkable problem-solving abilities and apparent self-awareness" (IBR, 2024). As reviewed in the Behavioral and Brain Sciences, "vertebrates, cephalopod mollusks, and euarthropods independently converged onto high levels of brain and cognitive complexity," and "this macroevolutionary trend was coupled with and facilitated by the acquisition of a small set of pivotal traits, used in visuomotor control of three-dimensional and targeted movements" (Coombs and Trestman, 2025).

The New York Declaration on Animal Consciousness, reflecting the assessment of 288 researchers, "suggests that consciousness is now a 'realistic' possibility not only in birds and mammals but also in reptiles, amphibians, fish, cephalopods like octopuses and cuttlefish, crustaceans like hermit crabs and crayfish, and insects like bees and fruit flies" (Özen et al., 2024). This declaration focuses on "phenomenal consciousness" or "sentience," reflecting the growing scientific consensus that subjective experience is not uniquely human but is likely widespread across the animal kingdom.

4.2 Convergent Neural Architectures for Consciousness

The observation that consciousness-related capacities have evolved independently in phylogenetically distant lineages with radically different neural organizations has profound implications for theories of consciousness. It suggests that consciousness is not dependent on any specific anatomical structure, such as the mammalian neocortex, but rather on certain abstract organizational principles that can be instantiated in diverse neural architectures. The identification of these principles is a central goal of comparative consciousness research.

A systematic application of neural markers of consciousness derived from human studies to animal data was conducted by Ehret and Romand (2022), who "analyzed published human data primarily on event-related potentials and brain-wave generation during perception and responding to sensory stimuli and extracted neural markers indicating that a person became aware or conscious of the perceived stimulus. These neural correlates of consciousness were then applied to sets of corresponding data from various animals including several species of mammals, and one species each of birds, fish, cephalopods, and insects." They found that "the neural markers from studies in humans could also successfully be applied to the mammal and bird data suggesting that species in these animal groups can become subjectively aware of and conscious about perceived stimuli" (Ehret & Romand, 2022). Notably, "fish, cephalopod and insect data remained inconclusive," suggesting either that these species may not possess the same forms of consciousness or that different neural markers may be needed to detect it.

The convergence across mammalian, avian, and potentially cephalopod lineages suggests that the following organizational principles may constitute the universal neural substrate of consciousness:

1. Hierarchical organization with recurrent connectivity (Zhang, 2012). Conscious processing requires not merely feedforward propagation of information through sensory hierarchies but sustained, bidirectional interactions between higher and lower processing levels. This recurrent connectivity enables the system to maintain and manipulate representations over extended time periods—a capacity that distinguishes conscious from unconscious processing.
2. High-dimensional integration. The capacity to integrate information across diverse processing modules (Zhang, 2016d) into unified representations is a hallmark of conscious-capable architectures. This integration must be accompanied by sufficient differentiation to support the vast repertoire of possible conscious contents. The balance between integration and differentiation may be optimized at a critical threshold in recurrent networks (Tagliazucchi, 2017; Zhang, 2012).
3. Thalamocortical (or analogous) resonance. In mammals, the thalamocortical system provides the oscillatory infrastructure that coordinates activity across distributed cortical regions. The intralaminar thalamic nuclei, in particular, serve as critical nodes for regulating the synchronization and desynchronization of cortical activity. Analogous structures may serve similar functions in birds (the thalamic-pallial loops) and potentially in cephalopods (central-peripheral integrative loops), though the evidence for this in invertebrates remains limited.
4. Neuromodulatory regulation of global brain state. The capacity to transition between global brain states—from wakefulness through sleep stages to REM sleep—appears to be a shared feature of mammals and birds, and may be present in more distant lineages. This regulation is achieved through diffuse neuromodulatory systems that adjust the excitability and signal-to-noise characteristics of neural networks on a global scale.
5. Long-range connectivity enabling global information integration. Conscious-capable architectures feature extensive long-range connections that enable information to be shared across distant processing modules. In mammals, these connections are mediated by the corpus callosum and long-range association fibers; in birds, by pallial commissures and association tracts; in cephalopods, by the connectivity between the central brain and the peripheral arm nervous system.

The observation that "gamma-wave generation for functional coupling of brain areas in aware/conscious states is energetically highly cost-intensive" (Ehret & Romand, 2022) suggests that consciousness may have evolved because it provides significant fitness advantages that offset its metabolic costs. Consciousness, in this view, emerges as an evolutionarily advantageous strategy for integrating information across specialized processing modules, enabling flexible, context-sensitive behavioral responses to complex environmental challenges.

5 Causal Intervention Studies: Anesthesia, Stimulation, and Manipulation

The strongest evidence for causal relationships (Zhang, 2021a-c) between neural activity and consciousness comes from intervention studies in which specific neural manipulations produce predictable changes in conscious state. The study of general anesthesia provides the most powerful and ethically tractable experimental paradigm for reversibly eliminating and restoring consciousness, allowing within-subject comparisons of the conscious and unconscious brain. Direct electrical stimulation of specific brain regions provides complementary evidence about the causal sufficiency of activity in particular structures for generating conscious experiences.

5.1 The Pharmacology of Unconsciousness

The remarkable fact that chemically diverse anesthetic agents—acting through completely different molecular mechanisms—all reliably induce loss of consciousness provides both a puzzle and an opportunity for consciousness science. Understanding how these diverse pharmacological interventions converge on a common behavioral endpoint can reveal the final common pathway through which consciousness is generated and lost.

Propofol, the most widely used intravenous anesthetic, acts primarily as a positive allosteric modulator of GABA_A receptors. By binding to specific β subunits, "propofol increases the receptor's sensitivity to GABA and prolongs channel opening, allowing a massive influx of negatively charged chloride ions. At higher therapeutic concentrations, propofol can directly activate the GABA_A receptor channel in the absence of GABA binding. This enhanced chloride conductance drives neuronal hyperpolarization, making the neuron significantly harder to excite" (Clarkson, 2026). The effects are subunit-specific: "targeting the β2 subunit mediates sedative and hypnotic effects, while the β3 subunit mediates the inhibition of responses to noxious stimuli. As a secondary mechanism, propofol also potentiates inhibitory glycine-gated chloride channels, particularly in the brainstem and spinal cord" (Clarkson, 2026). The functional consequence of these molecular actions is a global enhancement of inhibitory neurotransmission throughout the central nervous system, reducing cortical excitability and disrupting the recurrent processing loops necessary for consciousness.

The role of extra-synaptic GABA_A receptors in propofol's effects has been explored in neural population models. Hashemi et al. (2014) demonstrated that "the effect of the anesthetic drug propofol on GABAergic extra-synaptic receptors results in changes in neural population activity and the electroencephalogram (EEG). Increased tonic inhibition in inhibitory cortical neurons cause a dramatic increase in the power of both δ- and α-bands. Conversely, the effects of increased tonic inhibition in cortical excitatory neurons and thalamic relay neurons have the opposite effect and decrease the power in these bands. The increased δ-activity is in accord with observed data for deepening propofol anesthesia; but is absolutely dependent on the inclusion of extrasynaptic (tonic) GABA action in the model" (Hashemi et al., 2014).

Ketamine produces a fundamentally different form of unconsciousness—dissociative anesthesia, characterized by profound analgesia, altered consciousness, and relative preservation of muscle tone and respiration. At the molecular level, "ketamine blocks NMDA receptors, which triggers a glutamate surge, activating AMPA receptors. This inhibition of NMDA and stimulation of AMPA leads to the release of BDNF and downstream mTOR signaling" (Miller, Lopes, & McCurdy, 2025). At the circuit level, "ketamine is thought to act via NMDA receptors and HCN1 channels to produce brain oscillations that are related to these effects. At subanesthetic doses, ketamine preferentially blocks the NMDA receptors on GABAergic inhibitory interneurons, resulting in the disinhibition of downstream excitatory pyramidal neurons that is thought to facilitate increased gamma-band activity" (Tian et al., 2023). The dissociative state produced by ketamine represents a unique perturbation of consciousness—one in which subjective experience is profoundly altered rather than simply eliminated—providing a valuable window into the relationship between neural dynamics and the qualitative character of experience.

Critically, Schartner et al. (2017) demonstrated that "ketamine anesthesia paradoxically maintains or even increases the spatiotemporal complexity of cortical activity as measured by TMS-EEG, while psychedelic doses of ketamine are associated with increased spontaneous MEG signal diversity." This finding challenges the proposition that loss of consciousness is invariably accompanied by a reduction in neural complexity and suggests that different anesthetics may produce unconsciousness through fundamentally distinct network disruptions.

Convergence at the Systems Level.Despite their molecular diversity, recent evidence suggests that anesthetics may converge on a common systems-level mechanism. Eisen et al. (2026) demonstrated that "propofol, ketamine, and dexmedetomidine destabilize cortical dynamics to a similar degree. The destabilization was evident in neural responses to sensory inputs, which exhibited longer timescales during anesthesia." The researchers found that "all three anesthetics, despite their molecular differences, similarly affect cortical states by reducing dynamic stability. The estimated destabilization is corroborated by the slower recovery from sensory perturbations and longer stimulus-induced autocorrelation times observed during the anesthetic infusions. Destabilization is most prevalent in the low-frequency band of the population dynamics, linking it to the well-known increase in low-frequency power during anesthesia" (Eisen et al., 2026). These findings "suggest that cortical destabilization may be a shared neural correlate of anesthetic-induced unconsciousness and provide a plausible link to the low-frequency oscillations observed during anesthesia."

An important exception to the convergence pattern is that "decreased thalamic activity was associated with disconnectedness with all used anesthetics except for S-ketamine" (Kantonen et al., 2023). This finding highlights the unique pharmacological profile of ketamine and suggests that the thalamus may not be the sole bottleneck through which consciousness must pass. As summarized by Song (2025), key scientific questions remain about "how anesthetics, by acting on their pharmacological molecular targets to inhibit neuronal activity and interfere with or block information transmission, further lead to disorders of consciousness or loss of consciousness" and "what are the key nuclei, molecules, and their working mechanisms involved in the molecular neural mechanisms of regained consciousness after its loss."

5.2 Direct Brain Stimulation and the PCI

Direct electrical stimulation of the brain provides a powerful tool for testing causal hypotheses about the neural basis of consciousness. The ability of stimulation to elicit or disrupt specific conscious experiences provides strong evidence about which structures are causally sufficient for generating particular conscious contents.

Stimulation of the posterior cortical hot zone, particularly the temporo-parieto-occipital junction, can elicit complex, multimodal conscious experiences including visual hallucinations, auditory percepts, and out-of-body experiences. By contrast, stimulation of the prefrontal cortex typically does not elicit specific conscious experiences but may affect higher-order aspects of consciousness such as the sense of agency or the evaluation of perceptual decisions.

The PCI, as described in previous sections, has been validated as a reliable biomarker for consciousness across a wide range of conditions. Its clinical utility is demonstrated by studies such as that of Bai et al. (2016), who used TMS-EEG to evaluate the effects of repetitive TMS on consciousness in patients with disorders of consciousness, finding that "TMS-EEG might be an efficient assessment tool for evaluating rTMS protocol therapeutic efficiency in DOC." In this study, "the patient was diagnosed of a minimally conscious state minus (MCS-) by means of CRS-R at the interval of T0, however the TEP and PCI indicated the patient was vegetative state (VS)." Remarkably, "at the interval of T1, there was not any clinical behavioral improvement in CRS-R, but we could find significant changes in TEP, PCI, and GMFP. At the interval of T2 there was a significant increase of consciousness level according by CRS-R score, PCI value, TEP, and GMFP after 20 sessions of 10 Hz rTMS on the patient with DOC" (Bai et al., 2016).

5.3 Causal Hierarchy Model

Integrating evidence across molecular, cellular, and systems levels of analysis, I propose a causal hierarchy model of consciousness that distinguishes between different levels of causation:

Level 1: Molecular Prerequisites. The function of specific neurotransmitter receptors (GABA_A, NMDA, muscarinic acetylcholine receptors, etc.) and ion channels establishes the excitability and synaptic transmission properties of individual neurons. Anesthetics act at this level to alter the balance between excitation and inhibition. However, molecular events are neither necessary nor sufficient in themselves for consciousness—the same molecular targets can be modulated without affecting consciousness if the modulation is restricted to specific circuits or occurs at sub-anesthetic doses.

Level 2: Cellular and Local Circuit Events. The coordinated activity of neuronal populations within local circuits—including the balance between feedforward and feedback connections, the synchronization of oscillatory activity across cortical layers, and the generation of recurrent processing loops—constitutes the minimal neural event sufficient for a specific conscious content. This is the level at which the "neural correlate of consciousness" proper should be sought.

Level 3: Network Integration Events. The coordination of activity across multiple local circuits through long-range connections, thalamocortical loops, and inter-areal synchronization enables the binding of diverse features into unified conscious scenes and the global availability of conscious contents for behavioral control. This level corresponds to the transition from local phenomenal consciousness to global access consciousness.

Level 4: Global State Transitions. The regulation of the overall level of consciousness—the transition between wakefulness, sleep, anesthesia, and disorders of consciousness—depends on the integrity of the subcortical arousal systems and their interaction with the thalamocortical network. These global state transitions set the conditions under which the lower-level events can occur.

This hierarchical framework clarifies the relationship between different levels of analysis and resolves the apparent paradox of how diverse molecular interventions can converge on a common behavioral endpoint: they do so by disrupting the same intermediate-level processes—the recurrent processing loops and thalamocortical resonance—that constitute the core neural substrate of consciousness, even though they reach this substrate through different molecular routes.

6 Persistent Anomalies: Phenomena That Current Theories Cannot Explain

The systematic review of existing theories, anatomical constraints, comparative evidence, and causal intervention studies presented in the preceding sections reveals a set of robust empirical phenomena that resist satisfactory explanation by any single current theoretical framework. These anomalies are not peripheral curiosities but strike at the heart of the explanatory ambitions of consciousness science. They constitute the critical tests that any adequate theory must pass. I identify four central anomalies that serve as the logical starting point for theoretical innovation.

6.1 Anomaly 1: Phenomenal Consciousness Without Cognitive Access

Perhaps the most consequential anomaly in contemporary consciousness science is the apparent dissociation between phenomenal consciousness (subjective experience) and access consciousness (cognitive availability for report and behavioral control). The Overflow position, most prominently associated with Ned Block, holds that "phenomenal experience can occur without cognitive access—a case of consciousness that overflows report" (Usher et al., 2018). This position is supported by a substantial body of experimental evidence, including Sperling's classic partial-report paradigm and more recent studies using summary statistics and inattention paradigms.

Usher et al. (2018) presented data "supporting a mild Overflow position, according to which transient visual awareness can overflow report." They provided "new data showing that under data-limited conditions observers cannot discriminate a simple relation (same versus different) without discriminating the elements themselves and, based on additional computational considerations, we argue that this supports the interpretation that summary statistics (same/different) are grounded on the transient experience of elements." Furthermore, they examined "recent data from a variant of 'inattention blindness' and argue that contrary to widespread assumptions, it provides further support for Overflow by highlighting another factor, 'task relevance', which affects the ability to conceptualize and report (but not experience) visual elements" (Usher et al., 2018).

The most direct empirical demonstration of the P-without-A dissociation was provided by Amir et al. (2023), who "created a situation where participants lack online access to the stimulus yet are nevertheless able to retrospectively form judgements on its phenomenal, qualitative aspects." They further showed that "their performance cannot be fully explained by unconscious processing or by a response to stimulus offset," concluding that "P and A consciousness are not only conceptually distinct, but might also be teased apart empirically" (Amir et al., 2023). This finding "opens the gate to future studies pinpointing the neural correlates of the two types of consciousness" (Amir et al., 2023).

The implications of this anomaly are profound for theories of consciousness. The Global Neuronal Workspace theory, which identifies consciousness with global broadcasting and ignition in fronto-parietal networks, cannot easily accommodate the existence of phenomenal consciousness in the absence of such broadcasting. Proponents of GNW have argued that apparent cases of overflow are better explained by unconscious processing or by impoverished phenomenal experience (Cohen & Dennett, 2011; Dehaene, 2014). IIT, which identifies consciousness with integrated information (Φ) regardless of whether it is globally broadcast, can in principle accommodate overflow phenomena but struggles to explain why some information with high Φ is phenomenally conscious while other information with high Φ (as in the cerebellum, discussed below) is apparently not. Higher-order theories, which make consciousness dependent on meta-representations typically associated with prefrontal function, face the most direct challenge from overflow data.

6.2 Anomaly 2: The Cerebellum Puzzle

The cerebellum contains approximately 69 billion neurons—roughly 80% of the total neurons in the human brain—and is organized with an exquisitely regular, crystalline architecture optimized for precise temporal processing. Its parallel fiber system provides massive feedforward connectivity, and theoretical analyses suggest that it should have substantial information integration capacity. Yet complete cerebellar agenesis or extensive cerebellar lesions do not produce any apparent loss of conscious experience, and direct electrical stimulation of the cerebellum does not elicit conscious percepts. This stark dissociation between structural complexity and conscious capacity poses a fundamental challenge to any theory of consciousness.

The cerebellum puzzle is particularly acute for IIT. As Maillé and Lynn (2020) noted, IIT claims that it "can explain why the cortex is capable of producing conscious experience while the cerebellum is not, even though the cerebellum possesses up to four times more neurons." The IIT explanation is that "the paradoxes posed by cerebellum, the basal ganglia, and by other large neuronal aggregates that seem irrelevant for consciousness, are less disconcerting; their architecture is not suited for integrating information" (Tononi, 2012). Specifically, IIT argues that "when eight units are connected pairwise as parallel modules with minimal intermodular connectivity, the network specifies four separate, minimal cause-effect structures, each with very low Φ. Modular connectivities are found, for instance, in the cerebellum". Critics have challenged this resolution on the grounds that it relies too heavily on particular architectural assumptions and that the modularity of cerebellar circuitry does not necessarily preclude high Φ values under alternative modeling assumptions. Moreover, the claim that cerebellar architecture cannot support high Φ has been challenged by computational neuroscientists who point out that the cerebellum's extensive recurrent connections between deep cerebellar nuclei and the inferior olive, as well as its convergence zones, might under certain models generate substantial Φ.

The GNW theory explains the cerebellum puzzle by noting that cerebellar processing is not integrated into the global workspace—cerebellar outputs influence behavior through their projections to thalamocortical circuits, but the cerebellum itself lacks the long-range recurrent connectivity with prefrontal and parietal regions that characterizes the workspace. However, this explanation raises the question of why cerebellar outputs can influence behavior without the cerebellum itself being conscious, which in turn raises deeper questions about the relationship between consciousness and action control.

Recurrent processing theory faces the cerebellum puzzle as a challenge but can address it by emphasizing that "the feed-forward activation of cells in any area of the brain is not sufficient to generate awareness, but that recurrent processing, mediated by horizontal and feedback connections is necessary" (Lamme, 2000). The cerebellum's architecture, while highly structured and computationally powerful, is dominated by feedforward processing with relatively limited sustained recurrent loops of the type found in the cerebral cortex. The cerebellum's output pathways—primarily through the deep cerebellar nuclei to thalamus and then to cortex—mean that cerebellar computations can influence behavior without themselves having the kind of sustained, self-reinforcing recurrent dynamics that recurrent processing theories identify as the substrate of consciousness.

6.3 Anomaly 3: Convergent Anesthetic Mechanisms Without a Common Target

As described in the previous section, pharmacologically diverse anesthetics that act on completely different molecular targets—GABA_A receptors (propofol), NMDA receptors (ketamine), α2-adrenergic receptors (dexmedetomidine), and multiple other targets—all converge on loss of consciousness. The finding that "propofol, ketamine, and dexmedetomidine destabilize cortical dynamics to a similar degree" (Eisen et al., 2026) suggests that this convergence occurs at the systems level despite molecular diversity. However, the precise nature of this systems-level convergence remains debated, and the fact that it can be achieved through fundamentally different mechanisms challenges theories that identify consciousness too closely with any single neural process.

The finding that ketamine can maintain or increase neural complexity while producing unconsciousness is particularly problematic for theories that identify consciousness with complexity per se. The preservation of thalamic activity during ketamine anesthesia (in contrast to its suppression by propofol and dexmedetomidine) further suggests that there may be multiple routes to unconsciousness, only some of which involve thalamic deactivation (Kantonen et al., 2023). As the research indicates, "the mechanism of loss of consciousness (LOC) under anesthesia is unknown. Because consciousness depends on activity in the cortico-thalamic network, anesthetic actions on this network are likely critical for LOC" (Mashour & Hudetz, 2018), yet the diversity of anesthetic mechanisms suggests that this network can be disrupted at multiple nodes and through multiple mechanisms.

6.4 Anomaly 4: Convergent Evolution of Consciousness in Radically Different Neural Architectures

As described in the previous section, consciousness-related cognitive capacities have evolved independently in mammalian, avian, and likely cephalopod lineages, despite their radically different neural architectures. The mammalian neocortex (six-layered), the avian pallium (nuclear, non-layered), and the cephalopod nervous system (largely distributed with peripheral ganglia) share no direct structural homology at the level of gross anatomy, yet all appear to support forms of conscious processing.

This convergence poses a fundamental challenge to theories that identify consciousness with specific mammalian brain structures or organizational features. It demands that we identify the abstract organizational principles—rather than specific anatomical structures—that are necessary and sufficient for consciousness. As IBR (2024) argued, "among vertebrates, neural architectures distinct from primates' (e.g., cetaceans and corvids) enable advanced cognitive abilities; the distributed neurological networks of cephalopods, eg, octopuses, enable remarkable problem-solving abilities and apparent self-awareness," suggesting that "there are lots of recipes for consciousness, and they're all valid as long as they implement the right computational principles."

The existence of multiple evolutionary trajectories toward consciousness suggests that the phenomenon is not tied to a particular neural implementation but emerges from computational principles that can be realized in diverse physical substrates. This insight has profound implications for theory building: a successful theory of consciousness must be stated at a level of abstraction that captures the universal principles without being tied to the contingent details of any particular neural architecture.

7 The Feedback Integration Theory: A Novel Theoretical Synthesis

The systematic review of existing theories, anatomical constraints, dynamic signatures, comparative evidence, causal interventions, and persistent anomalies presented above reveals a landscape of partial insight and unresolved contradiction. Each major theory captures important aspects of the neural basis of consciousness—the global workspace theory correctly identifies the importance of widespread information sharing for access consciousness; IIT correctly identifies the importance of integration and differentiation for conscious experience; recurrent processing theory correctly identifies the temporal dynamics of conscious vs. unconscious processing—yet none provides a fully satisfactory account of the full range of empirical phenomena.

The Feedback Integration Theory (FIT) is proposed here as a novel theoretical synthesis that builds on the strengths of existing frameworks while resolving their most significant anomalies (Fig. 1). FIT is grounded in the convergent empirical finding that the neural signature most consistently associated with conscious experience across paradigms, modalities, and species is the recurrent, feedback-driven integration of information within hierarchically organized neural architectures.

7.1 Meta-Theoretical Position

FIT occupies a specific position in the theoretical landscape that might be characterized as functional-structural integrationism. The theory holds that consciousness is constitutively realized by a particular type of information processing operation—feedback integration—that requires specific structural features in the implementing neural architecture (Zhang, 2016d). The theory is thus functionalist in its emphasis on the computational operations that generate consciousness, but structuralist in its specification of the neural architectural features required to implement those operations. The theory is not a form of panpsychism: consciousness is not an intrinsic property of all matter or all information processing, but rather emerges only when specific architectural and dynamic conditions are met.

The theory is explicitly physicalistin its ontological commitments: consciousness is not an additional substance or force beyond the physical, but rather a natural phenomenon that arises from, and is constituted by, specific patterns of physical activity in complex neural systems. The appearance of an "explanatory gap" between physical processes and subjective experience reflects, in this view, the conceptual limitations of our current frameworks rather than any ontological dualism. The relationship between neural activity and conscious experience is one of constitution, not causation: the feedback-integrated activity of neural circuits simply is conscious experience, viewed from the first-person perspective.

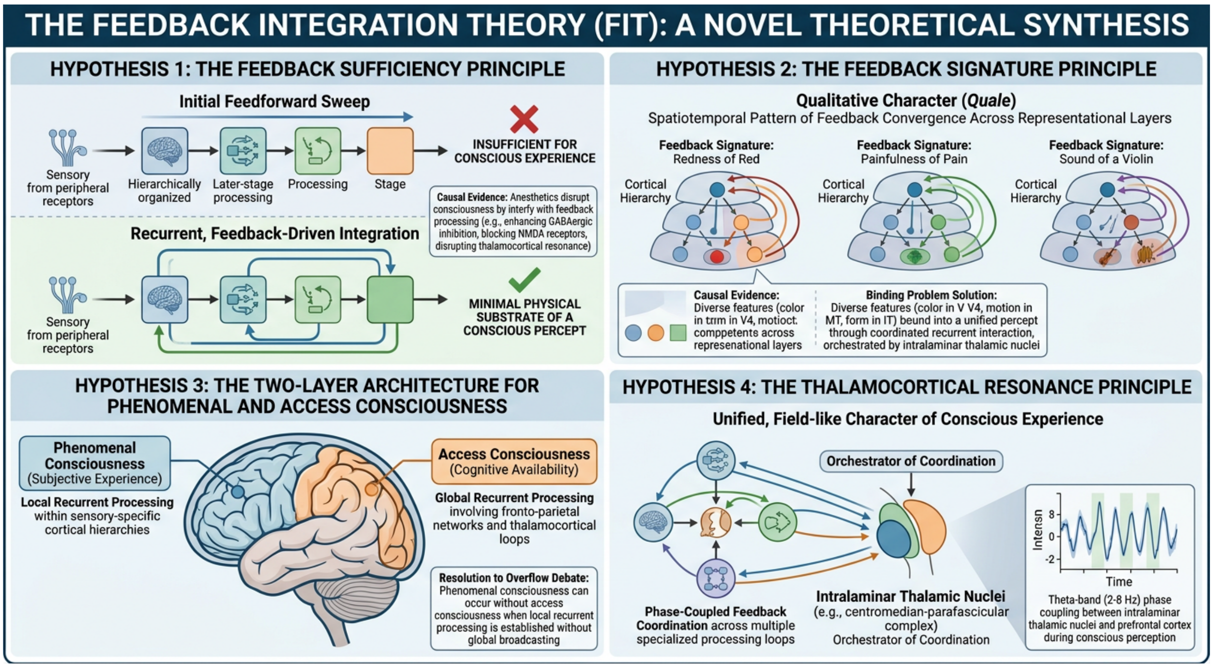

Fig. 1 The Feedback Integration Theory: A Novel Theoretical Synthesis.

7.2 Core Hypotheses and Principles

FIT is organized around four interconnected hypotheses, each addressing a different aspect of the consciousness phenomenon and each supported by converging lines of evidence from multiple experimental traditions.

Hypothesis 1: The Feedback Sufficiency Principle.The minimal physical substrate of a conscious percept is a recurrently connected neural loop capable of sustaining temporally structured feedback processing for at least one complete cycle of recurrent interaction. The initial feedforward sweep of activation—the rapid propagation of sensory information from peripheral receptors through hierarchically organized processing stages—is insufficient to generate conscious experience, regardless of its complexity, anatomical extent, or behavioral efficacy. Consciousness arises specifically when later-stage processing feeds back to modulate earlier-stage processing, creating a self-sustaining, temporally extended neural representation that integrates bottom-up sensory signals with top-down contextual information, predictions, and expectations.

This principle is directly supported by converging evidence from visual masking studies, binocular rivalry paradigms, and the neurophysiology of anesthesia. The finding that "the feed-forward activation of cells in any area of the brain is not sufficient to generate awareness, but that recurrent processing, mediated by horizontal and feedback connections is necessary" (Lamme, 2000) has been replicated across multiple experimental paradigms and constitutes one of the most robust findings in consciousness neuroscience. The fact that diverse anesthetics disrupt consciousness specifically by interfering with feedback processing—either by enhancing GABAergic inhibition that prevents recurrent excitation, or by blocking NMDA receptors that are critical for sustained recurrent activity, or by disrupting the thalamocortical resonance that coordinates feedback across distant regions—provides causal evidence for the feedback sufficiency principle.

Hypothesis 2: The Feedback Signature Principle.The specific qualitative character (quale) of a conscious experience is determined by the particular spatiotemporal pattern of feedback convergence across representational layers. Different conscious contents—the redness of red, the painfulness of pain, the sound of a violin—are distinguished not by which neurons are active per se, but by the specific pattern of recurrent interactions among neurons across the cortical hierarchy. A given conscious content corresponds to a unique "feedback signature": a characteristic pattern of top-down modulation that integrates information across multiple levels of representation.

This principle provides a natural explanation for the binding problem: diverse features processed in different cortical areas (color in V4, motion in MT, form in IT) are bound into a unified conscious percept because they are integrated through a common pattern of feedback coordination, rather than through convergence onto a single "consciousness neuron" or through synchronization at a particular frequency band. The binding is achieved through the coordinated recurrent interaction of the entire processing hierarchy, with the intralaminar thalamic nuclei serving as the orchestrator of this coordination.

Hypothesis 3: The Two-Layer Architecture for Phenomenal and Access Consciousness.The distinction between phenomenal consciousness (subjective experience) and access consciousness (cognitive availability) maps onto the neuroscientific distinction between two types of recurrent processing: local recurrent processing within sensory-specific cortical hierarchies generates phenomenal consciousness, while global recurrent processing involving fronto-parietal networks generates access consciousness. This two-layer architecture provides a principled resolution to the overflow debate: phenomenal consciousness can occur in the absence of access consciousness whenever local recurrent processing is established without subsequent global broadcasting.

This hypothesis builds directly on Lamme's distinction between local and global recurrent processing, but extends it by specifying the precise architectural conditions under which each type of recurrent processing can occur. Local recurrent processing requires intact feedback connections within a sensory-specific cortical hierarchy—conditions that can be met even when fronto-parietal networks are lesioned, explaining the preservation of conscious experience in prefrontal patients. Global recurrent processing requires the additional engagement of fronto-parietal networks and thalamocortical loops, enabling the information that is locally conscious to be globally broadcast and made available for report and behavioral control.

Hypothesis 4: The Thalamocortical Resonance Principle.The unified, field-like character of conscious experience—the sense that diverse contents are experienced within a single, coherent phenomenal field—emerges from phase-coupled feedback coordination across multiple specialized processing loops, orchestrated by thalamocortical resonance mechanisms centered on the intralaminar thalamic nuclei. These nuclei, particularly the centromedian-parafascicular complex, provide the oscillatory infrastructure that binds distributed neural activity into the temporally coherent patterns characteristic of conscious states.

This hypothesis provides a mechanistic account of the critical role of the intralaminar thalamic nuclei in consciousness. These nuclei are uniquely positioned to coordinate activity across distributed cortical regions because they receive convergent input from multiple cortical areas and b rainstem arousal systems and project diffusely to widespread cortical targets. The theta-band (2-8 Hz) phase coupling between intralaminar thalamic nuclei and prefrontal cortex during conscious perception provides direct electrophysiological evidence for this coordination mechanism (Cacciatore et al., 2025; Fang et al., 2025).

7.3 Explaining the Phenomena

Resolution of Anomaly 1 (Consciousness Without Access). FIT explains consciousness without access as cases in which local recurrent processing within sensory hierarchies is established—generating phenomenal consciousness—without the subsequent engagement of global recurrent processing through fronto-parietal networks that would make the content access-conscious. The transient nature of overflow phenomena, documented in multiple experiments (Usher et al., 2018; Amir et al., 2023), reflects the fact that local recurrent processing can be established and then rapidly dissipate, leaving a trace of phenomenal experience that can be retrospectively accessed even though online access was never established. This explanation captures the empirical data without requiring the problematic claim that consciousness can exist without any neural substrate for report—the local recurrent processing that generates the phenomenal experience is a real neural event, but it is one that does not trigger the global ignition that would make it immediately reportable.

Resolution of Anomaly 2 (The Cerebellum Puzzle). FIT explains the cerebellum's lack of reportable consciousness in terms of its architectural unsuitability for the type of sustained recurrent processing that generates conscious experience. While the cerebellum is extraordinarily rich in neurons and precise in its computational operations, its circuitry is dominated by feedforward processing: mossy fiber and climbing fiber inputs are processed through the granule cell layer and Purkinje cells, with output via the deep cerebellar nuclei, in a largely feedforward architecture with relatively limited long-range recurrent connectivity. The recurrent loops that do exist within the cerebellum—for example, the recurrent connections between deep cerebellar nuclei and the inferior olive—operate on timescales and with connectivity patterns that do not support the kind of sustained, content-specific feedback integration that FIT identifies as the substrate of consciousness. Cerebellar computations are thus unconscious not because they lack complexity or computational power, but because they lack the specific type of recurrent processing architecture that constitutes conscious experience.

This explanation is consistent with the observation that cerebellar function is essential for precise motor control, timing, and procedural learning—functions that depend on the cerebellum's unique feedforward processing architecture. The conscious experience of movement is generated not in the cerebellum itself but in the cortical circuits that receive and integrate cerebellar outputs, which do have the requisite recurrent architecture. The cerebellum is thus an unconscious processor whose outputs contribute to conscious experience through their effects on cortical recurrent circuits, but which does not itself generate conscious experience.

Resolution of Anomaly 3 (Anesthetic Convergence). FIT explains the convergence of diverse anesthetics on loss of consciousness in terms of their common interference with the feedback integration processes that constitute conscious experience. Propofol's enhancement of GABAergic inhibition disrupts the excitatory-inhibitory balance required for sustained recurrent activity; ketamine's blockade of NMDA receptors disrupts the glutamatergic transmission that is critical for maintaining recurrent excitatory loops; dexmedetomidine's suppression of noradrenergic tone reduces the global excitability of cortical circuits, shifting them away from the critical dynamical regime that supports feedback integration. Despite their molecular diversity, all these agents converge on disrupting the same fundamental process: the capacity of neural circuits to sustain temporally extended, feedback-integrated activity.

The paradoxical preservation of neural complexity under ketamine anesthesia can be explained within FIT by recognizing that global complexity per se is not a sufficient condition for consciousness; what matters is the specific type of feedback integration that generates coherent, content-specific representations. Ketamine may produce high-complexity dynamics that are disorganized and lack the coherent feedback signatures that characterize conscious experience—akin to a symphony orchestra in which each instrument is playing loudly but without coordination, resulting in noise rather than music.

Resolution of Anomaly 4 (Convergent Evolution). FIT provides a natural explanation for the convergent evolution of consciousness in lineages with radically different neural architectures, because it identifies consciousness not with any specific anatomical structure but with an abstract computational principle—feedback integration—that can be instantiated in diverse physical substrates. The mammalian neocortex, the avian pallium, and the cephalopod nervous system all support consciousness because they have all evolved architectures that implement feedback integration: hierarchical organization with extensive bidirectional connectivity, high-dimensional information integration, and recurrence at multiple spatial and temporal scales.

The fact that these architectures are structurally non-homologous but computationally convergent provides strong support for the claim that feedback integration is the universal computational principle underlying consciousness, rather than a contingent feature of mammalian cortical organization. This insight has profound implications for the search for consciousness in non-biological systems: any system, regardless of its physical substrate, that implements feedback integration with sufficient complexity and recurrence would necessarily be conscious.

7.4 Quantitative, Falsifiable Predictions

A scientific theory of consciousness must generate predictions that are sufficiently precise and specific to be falsified by empirical data. FIT generates the following suite of quantitative, testable predictions:

Prediction 1: Temporal Threshold for Consciousness. If FIT is correct, then the minimal duration of feedback processing required for consciousness should be measurable and should correspond to at least one complete cycle of recurrent interaction within the relevant processing hierarchy. For visual consciousness, this threshold should be approximately 100-150 ms after stimulus onset—the time required for the initial feedforward sweep to reach higher visual areas and for feedback to return to early visual cortex. Stimuli that are presented for shorter durations or that are masked before feedback can be established should not generate conscious experience, even if they produce robust feedforward activation. This prediction can be tested using visual masking paradigms with precise temporal control, combined with layer-specific recordings of feedforward and feedback activity.

Prediction 2: Feedback Disruption and Loss of Specific Conscious Contents. If the feedback signature principle is correct, then selectively disrupting feedback connections to a specific cortical region should eliminate the specific conscious content associated with that region without affecting other conscious contents or global arousal. For example, optogenetic suppression of feedback from higher visual areas to V4 should specifically eliminate conscious color experience (rendering the world phenomenally in grayscale) while leaving form, motion, and depth perception intact. The same manipulation of feedforward connections should produce different effects. This prediction can be tested in animal models using optogenetic tools with cell-type and pathway specificity.

Prediction 3: Φ and Feedback Co-Variation. If FIT is correct, measurements of integrated information (Φ) and measurements of feedback integration strength should be highly correlated across brain states, but the relationship should be asymmetric: high feedback integration should be necessary for high Φ (consciousness), but high Φ alone should not be sufficient if the feedback architecture is inappropriate. Specifically, modular networks with high local Φ but minimal feedback integration should not be conscious, while networks with the same Φ but rich feedback connectivity should be. This prediction can be tested using computational models of neural networks with parametrically varied architectures.

Prediction 4: Thalamic Theta Coherence and Conscious Access. If the thalamocortical resonance principle is correct, then theta-band (2-8 Hz) phase coherence between intralaminar thalamic nuclei and the relevant cortical regions should be a necessary condition for conscious access to sensory content. Pharmacological or optogenetic disruption of this theta coherence should prevent conscious access without affecting unconscious sensory processing, while enhancement of theta coherence should facilitate conscious access. This prediction can be tested using combined thalamic and cortical recordings in human patients or animal models.

Prediction 5: Cross-Species Neural Signature. If FIT correctly identifies feedback integration as the universal computational principle of consciousness, then species that exhibit behavioral indicators of consciousness—such as corvids, parrots, cetaceans, and cephalopods—should show electrophysiological signatures of feedback integration that are structurally analogous to those observed in conscious humans (e.g., sustained recurrent activity in sensory hierarchies, feedback-dependent enhancement of stimulus representations, and thalamocortical or analogous phase coupling). The absence of such signatures in a species that otherwise passes behavioral tests of consciousness would falsify the theory.

Prediction 6: Anesthetic Sensitivity and Feedback Architecture. If FIT is correct, the sensitivity of a neural system to anesthetic-induced loss of consciousness should be proportional to the extent to which that system's conscious processing depends on feedback integration. Systems that rely primarily on feedforward processing should be less affected by anesthetics at doses that eliminate consciousness. This prediction can be tested by comparing the effects of anesthetics on different types of neural processing in the same organism.

7.5 The Hard Problem: A Physicalist Account of Qualia

The "hard problem" of consciousness, as formulated by Chalmers (2010), is the question of why and how physical processes in the brain are accompanied by subjective, qualitative experience. Why does the integration of neural information feel like anything at all? FIT addresses this question from a principled physicalist perspective that attempts to dissolve rather than solve the explanatory gap.

FIT's position is that the apparent mysteriousness of the relationship between physical processes and subjective experience arises from a category error in our conceptual framework, not from any fundamental ontological gap. The feeling of pain, the redness of red, the sound of a violin—these qualitative experiences are not effects or products of neural activity that need to be explained by reference to something beyond the physical; they are, rather, what the feedback-integrated activity of specific neural circuits feels like from the inside. The "inside" perspective—the first-person point of view—is not an additional fact about the world beyond the physical facts; it is the intrinsic perspective of a physical system that is sufficiently complex to have a perspective at all.

This position can be clarified through the following chain of reasoning:

(1) Consciousness is not a function that the brain performs; it is a state that the brain is in. When we ask, "What is the function of consciousness?" we are asking the wrong question. Consciousness is not a tool that evolution designed to solve particular problems; it is the intrinsic manifestation of a particular type of neural organization. The evolutionary function of the neural architectures that support consciousness may be to enable flexible, context-sensitive behavioral control, but consciousness itself is not the function—it is what the implementation of that function is like from the inside.

(2) Qualitative experience is the way a feedback-integrated system registers its own state. When a neural circuit engages in sustained recurrent processing, it does not merely process information about the external world; it also, and inseparably, registers the effects of its own processing. This self-registration—the fact that the system's state includes information about its own processing—is the physical basis of what we call qualia. Different qualia correspond to different patterns of feedback integration, each of which constitutes a distinct way for the system to register its own state.

(3) The apparent explanatory gap reflects the fact that we are trying to explain the first-person perspective using only third-person concepts. The feedback-integrated activity of a neural circuit and the qualitative experience that it constitutes are not two different things that need to be connected by a causal or explanatory bridge; they are the same thing, viewed from two different perspectives. The neural activity viewed from the outside (by an external observer using neuroimaging or electrophysiology) is the third-person description; the same neural activity experienced from the inside (by the system itself) is the first-person experience. There is no additional fact to be explained beyond the physical facts.

(4) This position does not eliminate consciousness or reduce it to something non-conscious; rather, it expands our conception of the physical to include the intrinsic, first-person perspective as a fundamental aspect of certain types of physical organization. Just as nineteenth-century physicists had to expand their conception

of the physical to include electromagnetic fields—entities that were not reducible to mechanical particles but were nonetheless physical—so we may need to expand our conception of the physical to include experience as a fundamental aspect of feedback-integrated systems. This is not dualism; it is an enriched physicalism that recognizes the first-person perspective as a real, irreducible feature of certain physical systems.

This account of the hard problem is not a final solution; it is a promissory note for a future science of consciousness in which the relationship between physical processes and subjective experience is no more mysterious than the relationship between molecular motion and heat—a relationship that was once deeply puzzling but is now understood as one of constitution rather than causation. The challenge for future research is to develop experimental paradigms that can directly probe the relationship between feedback integration patterns and subjective experience, refining our understanding of exactly which physical properties correspond to which qualitative features.

8 Conclusion: Toward a Unified Science of Consciousness

The Feedback Integration Theory presented in this paper represents an attempt to synthesize the vast and diverse literature on the biological foundations of consciousness into a coherent, mechanistic framework that respects the full complexity of the empirical evidence while providing principled resolutions to the most persistent anomalies in the field. The theory's central claim—that consciousness is constitutively realized by the feedback-driven integration of information within hierarchically organized neural architectures—is grounded in convergent evidence from neuroanatomy, neurophysiology, comparative cognition, causal intervention studies, and clinical neurology.

FIT advances the science of consciousness in several distinct ways: (1) it provides a unified framework that reconciles the apparently competing claims of the major existing theories. The global neuronal workspace theory's emphasis on widespread broadcasting is captured by FIT's account of access consciousness; IIT's emphasis on information integration is captured by FIT's feedback signature principle; recurrent processing theory's emphasis on the temporal dynamics of conscious vs. unconscious processing is directly incorporated into FIT's core hypotheses. (2) FIT resolves anomalies that have resisted explanation by any single existing theory, including phenomenal consciousness without access, the cerebellum puzzle, the convergence of diverse anesthetics on loss of consciousness, and the convergent evolution of consciousness in phylogenetically distant lineages. (3) FIT generates a suite of precise, falsifiable predictions that can guide future experimental research, transforming the theory from a post hoc explanation into a prospective research program.

The implications of FIT extend beyond the neuroscience of consciousness to broader questions about the nature of mind, the possibility of machine consciousness, and the ethical treatment of non-human animals. If consciousness is constitutively realized by feedback integration, then any system—biological or artificial—that implements this computational principle with sufficient complexity would necessarily be conscious. This has profound ethical implications for the development of artificial intelligence and for our treatment of animal species whose neural architectures may support conscious experience.

The science of consciousness has made remarkable progress since Crick and Koch first proposed the search for neural correlates of consciousness as a tractable research program. We now have detailed anatomical maps of the structures necessary for consciousness, sophisticated measures of the dynamic signatures that distinguish conscious from unconscious states, and causal evidence linking specific neural processes to specific aspects of conscious experience. Yet the field remains fragmented, with competing theories that often talk past each other rather than engaging in direct, testable confrontation. The adversarial collaboration model pioneered by the COGITATE consortium represents an important step toward resolving these theoretical disputes through rigorous empirical testing, and the development of increasingly precise quantitative models of neural dynamics will enable more exacting tests of theoretical predictions in the coming years.

The ultimate goal of consciousness science is not merely to identify the neural correlates of consciousness—the structures and processes that accompany conscious experience—but to understand the fundamental principles that make some physical systems conscious and others not. The Feedback Integration Theory proposes that this principle is recurrent, feedback-driven integration: the capacity of a system not merely to process information in a feedforward manner but to sustain and modulate its own activity through temporally extended recurrent interactions. This principle is simple enough to be stated in computational terms that transcend any particular physical implementation, yet specific enough to generate quantitative, testable predictions. Whether it proves to be the correct principle remains to be determined by future research, but the systematic analysis presented in this paper suggests that it provides the most comprehensive and coherent account of the currently available evidence, and offers the most promising framework for advancing our understanding of the biological foundations of conscious experience.

References

Amir YZ, Assaf Y, Yovel Y, et al. 2023. Experiencing without knowing? Empirical evidence for phenomenal consciousness without access. Cognition, 238: 105529. https://www.sciencedirect.com/science/article/pii/S0010027723001634

Baars BJ. 1988. A Cognitive Theory of Consciousness. Cambridge University Press, Cambridge, UK. https://books.google.co.jp/books/about/A_Cognitive_Theory_of_Consciousness.html?id=3tpFdVe3dacC&redir_esc=y

Baars BJ. 2002. The conscious access hypothesis: origins and recent evidence. Trends in Cognitive Sciences, 6(1): 47-52. https://www.sciencedirect.com/science/article/pii/S1364661300018192?via%3Dihub

Baciadonna L, Cornero FM, et al. 2021. Convergent evolution of complex cognition: Insights from the field of avian cognition into the study of self-awareness. Learning & Behavior, 49: 9-22. https://link.springer.com/article/10.3758/s13420-020-00434-5

Bai Y, Xia XY, Kang JN, Yin XX, et al. 2016. Evaluating the effect of repetitive transcranial magnetic stimulation on disorders of consciousness by using TMS-EEG. Frontiers in Neuroscience, 10: 473. https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2016.00473/full

Birch J, Schnell AK, Clayton NS. 2020. Dimensions of animal consciousness. Trends in Cognitive Sciences, 24(10): 789-802. https://www.sciencedirect.com/science/article/pii/S1364661320301923?via%3Dihub

Block N. 1995. On a confusion about a function of consciousness. Behavioral and Brain Sciences, 18(2): 227-247. https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/on-a-confusion-about-a-function-of-consciousness/061422BF0C50C5FF00927F9B6E879413

Block N. 2007. Consciousness, accessibility, and the mesh between psychology and neuroscience. Behavioral and Brain Sciences, 30(5-6): 481-499. https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/consciousness-accessibility-and-the-mesh-between-psychology-and-neuroscience/63BF0B7E5DB1654ED680E160D413F3D0

Boly M, Massimini M, Tsuchiya N, et al. 2017. Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. Journal of Neuroscience, 37(40): 9603-9613. https://www.jneurosci.org/content/37/40/9603

Brown R, Lau H, LeDoux JE. 2019. Understanding the higher-order approach to consciousness. Trends in

Cognitive Sciences, 23(9): 754-768. https://www.sciencedirect.com/science/article/pii/S1364661319301615

Cacciatore M, Magnani FG, Barbadoro F, et al. 2025. Thalamus and consciousness: A systematic review on thalamic nuclei associated with consciousness. Frontiers in Neurology, 16: 1509668. https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2025.1509668/full

Casali AG, Gosseries O, Rosanova M, et al. 2013. A theoretically based index of consciousness independent of sensory processing and behavior. Science Translational Medicine, 5(198): 198ra105. https://www.science.org/doi/10.1126/scitranslmed.3006294

Chalmers DJ. 2010. Facing up to the problem of consciousness. In: The Character of Consciousness (Chalmers DJ, ed). 3-34, Oxford University Press, New York, USA. https://philpapers.org/rec/CHAFUT-5

Clarkson C. 2026. Propofol: cellular mechanisms of propofol-induced anesthesia. Tulane University School of Medicine. https://tmedweb.tulane.edu/pharmwiki/doku.php/propofol

Cogitate Consortium, Ferrante O, Gorska-Klimowska U, et al. 2025. Adversarial testing of global neuronal workspace and integrated information theories of consciousness. Nature, 642: 133-142. https://www.nature.com/articles/s41586-025-08888-1

Cohen MA, Dennett DC. 2011. Consciousness cannot be separated from function. Trends in Cognitive Sciences, 15(8): 358-364. https://www.sciencedirect.com/science/article/pii/S1364661311001252

Coombs S, Trestman M. 2025. Multi-trait convergent trends in the evolution of brains and cognition. Behavioral and Brain Sciences, 48: e98. https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/multitrait-convergent-trends-in-the-evolution-of-brains-and-cognition/3A0A5BD6EC7FA20CA113B98FA7D40BDF

Crick F, Koch C. 1990. Towards a neurobiological theory of consciousness. Seminars in the Neurosciences, 2: 263-275. https://www.psychologie.uni-heidelberg.de/ae/allg/mitarb/ms/crick90.pdf

Dehaene S, Changeux JP. 2011. Experimental and theoretical approaches to conscious processing. Neuron, 70(2): 200-227. https://www.sciencedirect.com/science/article/pii/S0896627311002583

Dehaene S, Changeux JP, Naccache L, et al. 2006. Conscious, preconscious, and subliminal processing: A testable taxonomy. Trends in Cognitive Sciences, 10(5): 204-211. https://www.sciencedirect.com/science/article/pii/S1364661306000799

Dehaene S, Naccache L, et al. 2001. Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. Cognition, 79(1-2): 1-37. https://www.sciencedirect.com/science/article/pii/S0010027700001232?via%3Dihub

Dehaene S, Sergent C, Changeux JP. 2003. A neuronal network model linking subjective reports and objective physiological data during conscious perception. Proceedings of the National Academy of Sciences, 100(14): 8520-8525. https://www.pnas.org/doi/10.1073/pnas.1332574100

Brian E, Hannah K. 2023. Harvard Ascending Arousal Network Atlas Version 2.0. DRYAD. https://datadryad.org/dataset/doi:10.5061/dryad.zw3r228d2

Ehret G, Romand R. 2022. Awareness and consciousness in humans and animals – neural and behavioral correlates in an evolutionary perspective. Frontiers in Systems Neuroscience, 16: 941534. https://www.frontiersin.org/journals/systems-neuroscience/articles/10.3389/fnsys.2022.941534/full

Eisen AJ, Bardon AG, Ballesteros JJ, et al. 2026. Similar destabilization of neural dynamics under different general anesthetics. Cell Reports, 45(3): 117048. https://www.sciencedirect.com/science/article/pii/S2211124726001269

Farisco M, Changeux JP. 2023. About the compatibility between the perturbational complexity index and the global neuronal workspace theory of consciousness. Neuroscience of Consciousness, 2023(1): niad016.

https://academic.oup.com/nc/article/2023/1/niad016/7202202

Fang ZP, Dang YY, Ping AA, et al. 2025. Human high-order thalamic nuclei gate conscious perception through the thalamofrontal loop. Science, 388(6742): eadr3675. https://www.science.org/doi/10.1126/science.adr3675

Gallup GG. 1970. Chimpanzees: self-recognition. Science, 167(3914): 86-87. https://www.science.org/doi/10.1126/science.167.3914.86

Gallup GG. 1977. Self-recognition in primates: a comparative approach to the bidirectional properties of consciousness. American Psychologist, 32(5): 329-338. https://www.semanticscholar.org/paper/Self-recognition-in-primates%3A-A-comparative-to-the-Gallup/708d77d2ab8ed4e2cf42973384a3898d82d5f6d3

Giacino JT, Fins JJ, Laureys S, et al. 2014. Disorders of consciousness after acquired brain injury: the state of the science. Nature Reviews Neurology, 10(2): 99-114. https://www.nature.com/articles/nrneurol.2013.279

Hohwy J. 2007. The search for neural correlates of consciousness. Philosophy Compass, 2: 461-474. https://compass.onlinelibrary.wiley.com/doi/abs/10.1111/j.1747-9991.2007.00086.x

Horn AKE, Němcová V, ten Donkelaar HJ, Overeem S. 2020. The reticular formation and the neuromodulatory systems. In: Clinical Neuroanatomy. 257-307, Springer. https://research.tue.nl/en/publications/the-reticular-formation-and-the-neuromodulatory-systems/

Hudetz AG, Mashour GA. 2016. Disconnecting consciousness: is there a common anesthetic end point? Anesthesia & Analgesia, 123(5): 1228-1240. https://journals.lww.com/anesthesia-analgesia/fulltext/2016/11000/disconnecting_consciousness__is_there_a_common.23.aspx

Hashemi M, Hutt A, Sleigh J. 2014. Anesthetic action on extra-synaptic receptors: effects in neural population models of EEG activity. Frontiers in Systems Neuroscience, 8: 232. https://www.frontiersin.org/journals/systems-neuroscience/articles/10.3389/fnsys.2014.00232

IBR. 2024. Neural diversity and the evolution of mind: expanding our understanding of consciousness. Institute for Biodefense Research. https://biodefenseresearch.org/neural-diversity-and-the-evolution-of-mind-expanding-our-understanding-of-consciousness/

Kantonen J, Laaksonen L, Alkire M, et al. 2023. Decreased thalamic activity is a correlate for disconnectedness during anesthesia with propofol, dexmedetomidine and sevoflurane but not S-ketamine. Journal of Neuroscience, 43(26): 4884-4895. https://www.jneurosci.org/content/43/26/4884

Koch C, Massimini M, Boly M, et al. 2016. Neural correlates of consciousness: progress and problems. Nature Reviews Neuroscience, 17(5): 307-321. https://www.nature.com/articles/nrn.2016.22

Lamme VAF. 2000. Neural mechanisms of visual awareness: A linking proposition. Brain and Mind, 1: 385-406. https://link.springer.com/article/10.1023/A:1011569019782

Lamme VAF. 2003. Why visual attention and awareness are different. Trends in Cognitive Sciences, 7(1): 12-18. https://www.sciencedirect.com/science/article/pii/S136466130200013X?via%3Dihub

Lamme VAF, Roelfsema PR. 2000. The distinct modes of vision offered by feedforward and recurrent processing. Trends in Neurosciences, 23(11): 571-579. https://www.sciencedirect.com/science/article/pii/S016622360001657X?via%3Dihub

Lamme VAF, Zipser K, Spekreijse H. 2002. Masking interrupts figure-ground signals in V1. Journal of Cognitive Neuroscience, 14(7): 1044-1053. https://direct.mit.edu/jocn/article-abstract/14/7/1044/3691/Masking-Interrupts-Figure-Ground-Signals-in-V1?redirectedFrom=fulltext

Lara AH, Wallis JD. 2015. The role of prefrontal cortex in working memory: A mini review. Frontiers in Systems Neuroscience, 9: 173. https://www.frontiersin.org/journals/systems-neuroscience/articles/10.3389/fnsys.2015.00173/full

Lau H, Rosenthal D. 2011. Empirical support for higher-order theories of conscious awareness. Trends in Cognitive Sciences, 15(8): 365-373. https://www.sciencedirect.com/science/article/pii/S1364661311001057?via%3Dihub

Maillé S, Lynn M. 2020. Reconciling current theories of consciousness. Journal of Neuroscience, 40(10): 1994-1996. https://www.jneurosci.org/node/648171.full.print

Martinaud O. 2018. Anatomical correlates of consciousness. Revue de Neuropsychologie, 10(4): 322-326. https://www.jle.com/en/revues/nrp/e-docs/anatomie_de_la_conscience_313405/article.phtml

Mashour GA, Hudetz AG. 2018. Neural correlates of unconsciousness in large-scale brain networks. Trends in Cognitive Sciences, 41(3): 150-160. https://www.sciencedirect.com/science/article/pii/S016622361830016X?via%3Dihub

Massimini M, Ferrarelli F, Huber R, et al. 2005. Breakdown of cortical effective connectivity during sleep. Science, 309(5744): 2228-2232. https://www.science.org/doi/10.1126/science.1117256

Michel M, Lau H. 2020. On the dangers of conflating strong and weak versions of a theory of consciousness. Phylosophy and the Mind Sciences, 1: II. https://philosophymindscience.org/index.php/phimisci/article/view/8952

Miller C, Lopes B, McCurdy A. 2025. The collective lie in ketamine therapy: A call to realign clinical practice with neurobiology. Frontiers in Psychiatry, 16: 1610335. https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2025.1610335/full

Mogensen J, Overgaard M. 2018. Reorganization of the connectivity between elementary functions as a common mechanism of phenomenal consciousness and working memory: from functions to strategies. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1755): 20170346. https://royalsocietypublishing.org/rstb/article/373/1755/20170346/30534/Reorganization-of-the-connectivity-between

Moruzzi G, Magoun HW. 1949. Brain stem reticular formation and activation of the EEG. Electroencephalography and Clinical Neurophysiology, 1(1-4): 455-473. https://www.sciencedirect.com/science/article/pii/0013469449902199

Malcom K. 2023. Geometry of the brain, dimensions of the mind. Michigan Medicine Health Lab. https://www.michiganmedicine.org/health-lab/geometry-brain-dimensions-mind

Oizumi M, Albantakis L, Tononi G. 2014. From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. PLoS Computational Biology, 10(5): e1003588. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003588

Özen S, Özlav E, Rona E. 2024. Joyful leaping dolphins, brooding crows... consciousness in animals. Kurious, Koç University. https://kurious.ku.edu.tr/en/joyful-leaping-dolphins-brooding-crows-consciousness-in-animals/

Pang KF. 2023. The many dimensions of consciousness. Psychology Today. https://www.psychologytoday.com/sg/blog/consciousness-and-beyond/202305/the-many-dimensions-of-consciousness

Parvizi J, Damasio A. 2001. Consciousness and the brainstem. Cognition, 79(1-2): 135-160. https://www.sciencedirect.com/science/article/pii/S001002770000127X

Plets C. 1998. Anatomical substrates of consciousness. European Journal of Anaesthesiology, 15: 4-5. https://journals.lww.com/ejanaesthesiology/fulltext/1998/01001/anatomical_substrates_of_consciousness.3

.aspx

Schartner MM, Carhart-Harris RL, Barrett AB, et al. 2017. Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin. Scientific Reports, 7: 46421. https://www.nature.com/articles/srep46421

Silvanto J, Lavie N, Walsh V. 2005. Double dissociation of V1 and V5/MT activity in visual awareness. Cerebral Cortex, 15(11): 1736-1741. https://academic.oup.com/cercor/article/15/11/1736/296944

Smythies J. 1997. The functional neuroanatomy of awareness: with a focus on the role of various anatomical systems in the control of intermodal attention. Consciousness and Cognition, 6(4): 455-481. https://www.sciencedirect.com/science/article/pii/S1053810097903159?via%3Dihub

Song XJ. 2025. Principles of shutdown and reboot of consciousness in the anesthetized brain and prospects for application. Science & Technology Review, 43(19): 23-29. https://castjournals.cast.org.cn/joweb/kjdb/CN/search?searchResult=1&keyword=Principles%20of%20shutdown%20and%20reboot%20of%20consciousness%20in%20the%20anesthetized%20brain%20and%20prospects%20for%20application

Stikvoort W,Pérez-Ordoyo E, Mindlin I, et al. 2025. Nonequilibrium brain dynamics elicited as the origin of perturbative complexity. PLoS Computational Biology, 21(6): e1013150. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1013150

Tagliazucchi E. 2017. The signatures of conscious access and its phenomenology are consistent with large-scale brain communication at criticality. Consciousness and Cognition, 55: 136-147. https://www.sciencedirect.com/science/article/abs/pii/S1053810017300478

Tapia E, Beck DM. 2014. Probing feedforward and feedback contributions to awareness with visual masking and transcranial magnetic stimulation. Frontiers in Psychology, 5: 1173. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2014.01173/full

Tian FY, Lewis LD, Zhou DW, et al. 2023. Characterizing brain dynamics during ketamine-induced dissociation and subsequent interactions with propofol using human intracranial neurophysiology. Nature Communications, 14: 1748. https://www.nature.com/articles/s41467-023-37463-3

Tong F, Nakayama K, Vaughan JT, et al. 1998. Binocular rivalry and visual awareness in human extrastriate cortex. Neuron, 21(4): 753-759. https://www.sciencedirect.com/science/article/pii/S0896627300805929

Tononi G. 2004. An information integration theory of consciousness. BMC Neuroscience, 5: 42. https://link.springer.com/article/10.1186/1471-2202-5-42

Tononi G. 2008. Consciousness as integrated information: a provisional manifesto. The Biological Bulletin, 215(3): 216-242. https://www.journals.uchicago.edu/doi/full/10.2307/25470707

Tononi G. 2012. Integrated information theory of consciousness: An updated account. Archives Italiennes de Biologie, 150: 290-326. http://www.esalq.usp.br/lepse/imgs/conteudo_thumb/Integrated-information-theory-of-consciousness-an-updated-account.pdf

Tononi G, Boly M, Massimini M, et al. 2016. Integrated information theory: from consciousness to its physical substrate. Nature Reviews Neuroscience, 17: 450-461. https://www.nature.com/articles/nrn.2016.44

Tononi G, Koch C. 2015. Consciousness: here, there and everywhere? Philosophical Transactions of the Royal Society B: Biological Sciences, 370(1668): 20140167. https://royalsocietypublishing.org/rstb/article/370/1668/20140167/22537/Consciousness-here-there-and-everywhere

Usher M, Bronfman ZZ, Talmor S, et al. 2018. Consciousness without report: Insights from summary statistics and inattention 'blindness'. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1755): 20170354. https://royalsocietypublishing.org/rstb/article/373/1755/20170354/30457/Consciousness-without-report-insights-from-summary

Vogt BA, Laureys S. 2005. Posterior cingulate, precuneal and retrosplenial cortices: Cytology and components of the neural network correlates of consciousness. Progress in Brain Research, 150: 205-217. https://www.sciencedirect.com/science/chapter/bookseries/abs/pii/S0079612305500153?via%3Dihub

Yang Y, et al. 2025. Chinese Neurosurgical Journal study provides insights into the role of the brainstem in disorders of consciousness. Chinese Neurosurgical Journal, 11: 42. https://sciencesources.eurekalert.org/news-releases/1109283

Zeman A. 2008. Consciousness: concepts, neurobiology, terminology of impairments, theoretical models and philosophical background. In: Handbook of Clinical Neurology, 90: 3-31. https://www.sciencedirect.com/science/chapter/handbook/pii/S0072975207017010

Zhang WJ. 2010. Computational Ecology: Artificial Neural Networks and Their Applications. World Scientific, Singapore. https://doi.org/10.1142/7436

Zhang WJ. 2012. Computational Ecology: Graphs, Networks and Agent-based Modeling. World Scientific, Singapore. https://doi.org/10.1142/8118

Zhang WJ. 2016a. A method for identifying hierarchical sub-networks / modules and weighting network links based on their similarity in sub-network / module affiliation. Network Pharmacology, 1(2): 54-65. http://www.iaees.org/publications/journals/np/articles/2016-1(2)/identifying-hierarchical-sub-networks-modules.pdf

Zhang WJ. 2016b. Detecting connectedness of network: A Matlab program and application in tumor pathways and a phylogenic network. Selforganizology, 3(4): 117-120. http://www.iaees.org/publications/journals/selforganizology/articles/2016-3(4)/detecting-connectedness-of-network.pdf

Zhang WJ. 2016c. How to find cut nodes and bridges in the network? A Matlab program and application in tumor pathways. Network Pharmacology, 1(3): 82-85. http://www.iaees.org/publications/journals/np/articles/2016-1(3)/find-cut-nodes-and-bridges-in-the-network.pdf

Zhang WJ. 2016d. Network chemistry, network toxicology, network informatics, and network behavioristics: A scientific outline. Network Biology, 6(1): 37-39

Zhang WJ. 2016e. Selforganizology: The Science of Self-Organization. World Scientific, Singapore. https://doi.org/10.1142/9685

Zhang WJ. 2017. Finding fundamental circuits in the network: A Matlab program and application in tumor pathway. Selforganizology, 4(1): 14-17. http://www.iaees.org/publications/journals/selforganizology/articles/2017-4(1)/finding-fundamental-circuits-in-the-network.pdf

Zhang WJ. 2018. Fundamentals of Network Biology. World Scientific Europe, London, UK. https://doi.org/10.1142/q0149

Zhang WJ. 2021a. A statistical simulation method for causality inference of Boolean variables. Network Biology, 11(4): 263-273. http://www.iaees.org/publications/journals/nb/articles/2021-11(4)/a-method-for-causality-inference-of-Boolean-variables.pdf

Zhang WJ. 2021b. Causality inference of linearly correlated variables: The statistical simulation and regression method. Computational Ecology and Software, 11(4): 154-161.

http://www.iaees.org/publications/journals/ces/articles/2021-11(4)/causality-inference-of-linearly-correlated-variables.pdf

Zhang WJ. 2021c. Causality inference of nominal variables: A statistical simulation method. Computational Ecology and Software, 11(4): 142-153. http://www.iaees.org/publications/journals/ces/articles/2021-11(4)/causality-inference-of-nominal-variables-with-statistical-simulation-method.pdf

Zhang WJ. 2027. Can generative AI produce consciousness? A cross-disciplinary literature review and critical analysis. Selforganizology, 14(3-4): 34-46. http://www.iaees.org/publications/journals/selforganizology/articles/2027-14(3-4)/can-generative-AI-produce-consciousness.pdf