JNφphi

The Neurobiology of Cognition in the Age of Artificial Intelligence: From Synaptic Plasticity to Cognitive Mapping

Taruna Ikrar*, Alfi Sophian

Abstract

Cognition emerges from complex interactions among molecular, cellular, and network-level processes in the brain that allow adaptive representation, reasoning, and behavioral regulation. Recent advances in neurobiology, combined with artificial intelligence (AI) and computational modeling, have illuminated previously inaccessible aspects of cognitive mechanisms—ranging from synaptic plasticity to large-scale cognitive mapping. This review explores how neural substrates underpin core cognitive processes such as attention, memory, and prediction, and how these biological architectures inspire AI systems through neuro-symbolic and neuromorphic approaches. We examine the dynamic relationship between predictive coding, hierarchical cortical networks, and consciousness as a distributed emergent property. Furthermore, new insights from connectomics, optogenetics, and neural decoding provide unprecedented clarity about the biophysical basis of cognition. The convergence between neurobiology and AI offers not only models of intelligence but also novel frameworks to understand self-referential cognition, agency, and ethical implications of synthetic minds. Ultimately, the neurobiology of cognition is entering a transformative era where understanding the brain's logic informs both human enhancement and machine consciousness.

Key Words: cognition, neurobiology, predictive coding, neuro-AI convergence, synaptic plasticity

1. Introduction

The study of cognition, which explores how the brain perceives, interprets, and acts upon information, has experienced a profound evolution from its philosophical origins to a mechanistic discipline grounded in neurobiology. Early explorations of the mind were largely descriptive and speculative, seeking to understand thought, perception, and behavior through introspection and logic rather than

Corresponding author: Taruna Ikrar
Address: Indonesia FDA, Jl. Percetakan Negara, No.23, Jakarta Pusat, 10560, Indonesia
e-mail ✉ taruna.ikrar@pom.go.id; alfi.sophian@pom.go.id

biological evidence. However, advances in neuroscience during the twentieth century transformed cognition from an abstract concept into an empirically measurable process. Contemporary neurobiology now defines cognition as the dynamic interaction of molecular signaling, neuronal circuitry, and coordinated brain networks that together enable adaptive behavior and complex thought (Churchland, 1989; Friston, 2010). This transformation represents a paradigm shift, positioning cognition as a product of biological computation rather than immaterial reasoning.

Historically, the cognitive revolution in the mid-twentieth century marked the beginning of this transition. Scholars such as George Miller and Herbert Simon conceptualized the mind as an information-processing system that encodes, stores, and retrieves knowledge (Miller, 2003; Newell & Simon, 1976). Yet, these early computational models remained symbolic, operating at the level of mental representations rather than their neurobiological basis. The emergence of cognitive neuroscience in the 1980s bridged this gap by combining cognitive theory with direct observations of brain activity. Techniques such as electrophysiology, positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) revealed specific neural correlates of attention, memory, and language (Gazzaniga, Ivry, & Mangun, 2019). These findings demonstrated that mental processes are instantiated in physical structures, leading to a materialist understanding of cognition as a property of neural systems rather than an abstract mental faculty.

195

At the cellular level, cognition is grounded in synaptic plasticity, the capacity of neurons to strengthen or weaken their connections in response to experience. This plasticity provides the biological foundation for learning and memory (Bliss & Collingridge, 1993). Molecular mechanisms such as calcium influx, NMDA receptor activation, and protein kinase signaling modulate the efficacy of synaptic transmission, thereby encoding experience within neural circuits (Kandel, 2013). Neuromodulators such as dopamine and serotonin further influence cognitive processes by regulating reward, motivation, and emotional tone (Schultz, 2015). These biochemical interactions illustrate how cognition arises from the collective behavior of countless molecular and cellular processes that continually reorganize to optimize adaptive function.

Over time, research has shown that cognition cannot be localized to single brain regions but rather emerges from distributed network activity. The human brain operates as a complex network in which specialized regions communicate through long-range connections to support perception, reasoning, and decision making (Sporns, 2018). Functional neuroimaging studies have revealed that cognitive tasks depend on dynamic coordination between the prefrontal cortex, parietal cortex, hippocampus, and other associative regions (Buschman & Miller, 2007). These interactions exhibit oscillatory

synchronization across multiple frequency bands, suggesting that temporal coherence among neuronal populations underlies the integration of information into coherent experience (Singer, 1999). Consequently, cognition is best understood as an emergent property of large-scale network dynamics rather than the activity of isolated modules.

A significant refinement of this network perspective comes from the growing appreciation of non-neuronal contributions to cognition. Glial cells, traditionally viewed as passive support structures, play active roles in maintaining neural homeostasis and modulating information flow. Astrocytes regulate synaptic transmission through neurotransmitter uptake and calcium signaling, while microglia refine neural circuits through synaptic pruning and immune surveillance (Fields, Woo, & Basser, 2015; Paolicelli et al., 2011). These discoveries have expanded the neurobiological model of cognition to encompass cellular diversity and intercellular communication, underscoring that the brain functions as an integrated ecosystem rather than a collection of independent neurons.

The predictive coding framework provides a unifying principle that connects molecular, cellular, and network-level processes of cognition. According to this theory, the brain continuously generates internal predictions about sensory input and adjusts its models based on the discrepancy between expected and actual signals (Friston, 2010). Cognition thus becomes an inferential process driven by the minimization of prediction error. This model explains a wide range of phenomena, from basic sensory perception to abstract reasoning and social understanding (Clark, 2013). Empirical studies demonstrate that cortical hierarchies are organized to transmit both top-down predictions and bottom-up error signals (Rao & Ballard, 1999). By constantly updating these internal models, the brain maintains perceptual stability and anticipatory control of behavior.

This predictive mechanism extends beyond perception to encompass motor planning, memory, and social cognition. In motor control, predictive coding allows the brain to anticipate the sensory consequences of movement, enabling fluid and precise action (Wolpert & Kawato, 1998). In memory and imagination, the same generative mechanisms allow the simulation of past or future events, facilitating planning and decision making. Even in social interaction, individuals rely on predictive models to infer others' intentions and emotions (Frith & Frith, 2012). By framing cognition as an ongoing process of model updating, predictive coding integrates perception, action, and learning into a single theoretical framework that unites neurobiology with computational principles.

The development of connectomics has further advanced the understanding of cognitive architecture. Large-scale initiatives such as the Human Connectome Project have mapped the intricate structural and functional pathways that support cognition (Van Essen

et al., 2013). This research reveals that brain organization follows a small-world topology characterized by high local clustering and short global path lengths, optimizing both specialization and integration (Bullmore & Sporns, 2009). Cognitive efficiency depends on the balance between these properties: too much segregation limits integration, while excessive connectivity leads to interference. Disorders such as schizophrenia, depression, and Alzheimer's disease can thus be viewed as network dysfunctions in which connectivity patterns are disrupted, impairing the coordination required for coherent thought and memory.

In parallel with these neuroscientific discoveries, artificial intelligence has emerged as a technological counterpart to biological cognition. Early AI research, inspired by symbolic logic, sought to emulate reasoning without replicating the brain's biological substrate (Newell & Simon, 1976). However, the resurgence of neural networks and deep learning has shifted focus toward biologically inspired architectures that mirror the hierarchical and distributed nature of the cortex (LeCun, Bengio, & Hinton, 2015). These artificial systems excel in pattern recognition, language processing, and decision optimization, offering new tools for modeling cognitive processes. Nevertheless, their limitations—such as lack of contextual understanding, emotional intelligence, and self-awareness—highlight the gap between artificial and biological cognition.

197

The emerging field of NeuroAI seeks to bridge this divide by integrating principles from neuroscience into machine learning and, conversely, using AI to decode the brain's computational patterns (Hassabis, Kumaran, Summerfield, & Botvinick, 2023). For example, spiking neural networks mimic the temporal coding used by biological neurons, while neuromorphic chips replicate the efficiency of synaptic computation (Roy, Jaiswal, & Panda, 2019). At the same time, AI algorithms are increasingly applied to brain imaging data to uncover latent patterns of connectivity and neural representation (Kell et al., 2018). This reciprocal exchange is reshaping both fields, suggesting that the future of cognitive science lies in the intersection between biology and computation.

The integration of neurobiology and AI also provokes deep philosophical and ethical questions. Can machines that emulate neural computation ever achieve genuine understanding or consciousness? Searle's (1980) Chinese Room argument posits that symbol manipulation alone cannot produce true comprehension, while Churchland (2007) contends that functional equivalence at the neural level could, in principle, yield conscious experience. These debates emphasize that cognition involves more than information processing; it encompasses embodiment, affect, and subjective awareness. As AI systems increasingly replicate aspects of human cognition, society must confront new challenges related to autonomy, identity, and moral responsibility (Yuste et al., 2017).

In medicine, insights into the neurobiology of cognition are transforming the diagnosis and treatment of neurological and psychiatric disorders. Cognitive impairment in conditions such as Alzheimer's disease, schizophrenia, and traumatic brain injury is increasingly understood as a failure of plasticity or network integration rather than isolated regional damage (Edison, Holland, & Minhas, 2022). Advances in neuroimaging, genomics, and computational modeling have made it possible to identify biomarkers of cognitive dysfunction and develop personalized interventions. Moreover, neurotechnologies such as deep brain stimulation, neurofeedback, and noninvasive brain stimulation offer new avenues for cognitive enhancement and rehabilitation. These developments highlight the clinical relevance of understanding cognition as a dynamic, biologically grounded process.

Taken together, these scientific and technological advances reveal cognition as a multiscale phenomenon that links molecular signaling, neural architecture, and behavioral adaptation. The brain does not merely process information; it constructs meaning through recursive, predictive, and embodied interactions with the environment. In this framework, cognition becomes both a biological function and an epistemic process that shapes how organisms understand and engage with the world. The convergence of neuroscience and artificial intelligence promises not only to deepen our comprehension of natural cognition but also to inspire novel forms of synthetic intelligence capable of emulating its flexibility and creativity.

2. Cellular and Molecular Basis of Cognition

Cognition at its core represents the brain's capacity to perceive, interpret, and respond to environmental stimuli through adaptive mechanisms that integrate molecular, cellular, and network-level processes. Understanding cognition from a neurobiological perspective therefore demands an exploration of the microscopic architecture of the nervous system—how molecular signaling and cellular plasticity translate into emergent cognitive functions such as learning, memory, and reasoning. Over the past three decades, breakthroughs in molecular neurobiology, neuroimaging, and electrophysiology have enabled the dissection of these processes at an unprecedented level of detail. This section provides a comprehensive review of the cellular and molecular underpinnings of cognition, focusing on synaptic plasticity, neurotransmission, intracellular signaling, neurotrophic regulation, and glial contributions, along with their integrative roles in higher-order cognitive processes.

At the molecular level, synaptic plasticity remains the cornerstone of cognitive function. First characterized in the hippocampus, long-term potentiation (LTP) and long-term depression (LTD) serve as

physiological correlates of learning and memory (Bliss & Collingridge, 1993). LTP describes the strengthening of synaptic transmission following repeated stimulation, while LTD involves a persistent weakening of synaptic efficacy after low-frequency activation. These bidirectional forms of plasticity depend on calcium influx through N-methyl-D-aspartate (NMDA) receptors, which act as molecular coincidence detectors for pre- and postsynaptic activity (Malenka & Bear, 2004). When glutamate binds to NMDA receptors coincident with postsynaptic depolarization, the magnesium block is relieved, allowing calcium ions to enter the neuron and trigger downstream signaling cascades that modulate synaptic strength.

The molecular machinery governing LTP and LTD involves a sophisticated interplay between kinases, phosphatases, and cytoskeletal remodeling proteins. Calcium/calmodulin-dependent protein kinase II (CaMKII) plays a pivotal role in LTP induction by phosphorylating α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor subunits and promoting their insertion into the postsynaptic membrane (Lisman et al., 2012). Conversely, LTD often recruits protein phosphatases such as calcineurin and protein phosphatase 1 to induce AMPA receptor internalization, thereby reducing synaptic efficacy (Collingridge et al., 2010). The dynamic regulation of receptor trafficking thus underlies the brain's capacity for experience-dependent modification, a property essential for encoding cognitive representations.

Beyond receptor signaling, structural plasticity of dendritic spines—the small protrusions on neuronal dendrites that serve as postsynaptic sites—is a morphological correlate of learning. Spine enlargement during LTP is accompanied by actin cytoskeleton reorganization driven by Rho-family GTPases, cofilin, and LIM kinase (Hotulainen & Hoogenraad, 2010). These structural changes are stabilized by brain-derived neurotrophic factor (BDNF), which activates the TrkB receptor and downstream MAPK and PI3K-Akt signaling pathways, reinforcing synaptic connections (Park & Poo, 2013). Deficits in BDNF signaling have been implicated in cognitive disorders, including depression, Alzheimer's disease, and schizophrenia (Autry & Monteggia, 2012), underscoring the tight coupling between molecular homeostasis and cognitive integrity.

Neurotransmitter systems provide another molecular substrate for cognition. Glutamate, the principal excitatory neurotransmitter in the brain, facilitates most forms of fast synaptic transmission and underlies both LTP and LTD (Watkins & Jane, 2006). GABAergic inhibition, mediated by γ-aminobutyric acid (GABA) receptors, ensures network stability by balancing excitation and preventing runaway activity that could impair information processing (Isaacson & Scanziani, 2011). Dopaminergic, serotonergic, and cholinergic systems further modulate cognitive states by regulating attention, motivation, and executive control. Dopamine, released from midbrain

structures such as the ventral tegmental area, gates synaptic plasticity in prefrontal and striatal circuits, thereby influencing working memory and decision-making (Seamans & Yang, 2004; Schultz, 2016). The cholinergic system, originating from the basal forebrain, enhances cortical signal-to-noise ratio and promotes learning by increasing neuronal responsiveness to inputs (Hasselmo & Sarter, 2011). Dysregulation of these modulatory systems contributes to cognitive impairments observed in disorders like Parkinson's disease, attention deficit hyperactivity disorder (ADHD), and dementia (Millan et al., 2012).

Recent evidence also emphasizes the crucial role of intracellular signaling and gene expression in sustaining cognitive processes over longer timescales. The consolidation of long-term memories requires the activation of transcription factors such as CREB (cAMP response element-binding protein), which regulates the expression of plasticity-related genes, including BDNF, Arc, and c-fos (Kida & Serita, 2014). These genes contribute to synaptic remodeling by promoting protein synthesis and cytoskeletal stability, thereby converting transient electrical activity into lasting structural changes. Epigenetic modifications, such as histone acetylation and DNA methylation, further modulate gene accessibility in response to experience (Day & Sweatt, 2011). Such mechanisms bridge the molecular and behavioral domains, suggesting that cognition is not merely an emergent property of network dynamics but also a product of molecular memory encoded within the genome.

200

In addition to neurons, glial cells—astrocytes, microglia, and oligodendrocytes—play indispensable roles in cognitive processing. Once considered passive support cells, astrocytes are now recognized as active participants in synaptic transmission, forming the so-called "tripartite synapse" (Araque et al., 2014). They release gliotransmitters such as ATP, D-serine, and glutamate, modulating synaptic efficacy and network oscillations that underlie attention and learning (Perea et al., 2009). Microglia, the brain's resident immune cells, shape synaptic architecture through activity-dependent pruning, eliminating weaker synapses while preserving functionally relevant ones (Schafer et al., 2012). Dysregulated microglial activity has been linked to neurodevelopmental and neurodegenerative disorders, highlighting their importance in maintaining cognitive homeostasis. Oligodendrocytes, through myelination of axons, regulate conduction velocity and temporal coordination among distributed neural assemblies—factors critical for synchronous activity and cognitive integration (Fields, 2015).

A growing body of work also implicates energy metabolism and mitochondrial dynamics in cognitive performance. Synaptic transmission is energetically demanding, with ATP consumption rates among the highest in the body (Harris et al., 2012). Mitochondria localize near synapses to supply ATP for neurotransmitter release and

to buffer calcium transients during high-frequency activity. Impaired mitochondrial function leads to synaptic deficits and cognitive decline, as observed in aging and neurodegenerative conditions (Kann & Kovács, 2007). Furthermore, reactive oxygen species (ROS), once thought to be purely deleterious, serve as signaling molecules regulating synaptic plasticity under physiological conditions (Massaad & Klann, 2011). Thus, bioenergetic balance and redox homeostasis form essential molecular foundations for cognitive function.

At the network level, the cellular mechanisms described above integrate to support distributed cognitive operations. The hippocampus, prefrontal cortex, and parietal regions form interconnected loops that encode spatial, episodic, and executive components of cognition. Synaptic plasticity within these circuits allows the brain to represent temporal sequences and causal relationships—a fundamental aspect of reasoning and prediction. For instance, hippocampal place cells and entorhinal grid cells create spatial cognitive maps, while prefrontal neurons encode rule-based associations necessary for abstract thought (O'Keefe & Nadel, 1978; Eichenbaum, 2017). These specialized cell assemblies operate within large-scale oscillatory frameworks coordinated by theta and gamma rhythms, which synchronize information flow across cortical and subcortical regions (Buzsáki & Draguhn, 2004). At the cellular level, such coordination depends on GABAergic interneurons and gap junction-mediated synchronization among pyramidal neurons.

Figure 1 (below) illustrates a simplified model of synaptic plasticity as the cellular basis of cognition, depicting key molecular actors—NMDA receptor activation, CaMKII signaling, and BDNF-mediated spine stabilization. The figure highlights how neurotransmitter release and receptor dynamics translate into structural and functional modifications that encode memory traces.

Recent advances in single-cell transcriptomics and proteomics have revealed molecular diversity among neurons and glia that correlates with cognitive specialization. Distinct gene expression profiles in cortical pyramidal neurons, inhibitory interneurons, and astrocytic subtypes suggest that cognition arises from the collective behavior of molecularly heterogeneous cell populations (Zeisel et al., 2018). Such heterogeneity allows for region-specific modulation of synaptic function, metabolic capacity, and plasticity thresholds—factors that together define the computational architecture of the brain. Moreover, new tools such as optogenetics and CRISPR-based gene editing enable causal manipulation of molecular pathways implicated in cognition, offering experimental precision to dissect mechanisms once considered inaccessible (Deisseroth, 2015).
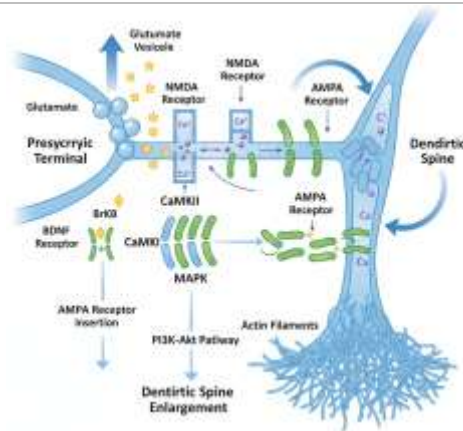
Figure 1. Schematic representation of the molecular and cellular mechanisms underlying synaptic plasticity in cognitive processes. High-frequency stimulation induces NMDA receptor-mediated Ca2+ influx, activating CaMKII and MAPK pathways that enhance AMPA receptor insertion and dendritic spine enlargement. BDNF released from presynaptic terminals binds to TrkB receptors, promoting actin polymerization and synaptic stabilization. These combined processes form the structural basis for long-term memory and learning.

From a systems perspective, the molecular basis of cognition cannot be isolated from its developmental and evolutionary context. During neurodevelopment, gradients of morphogens and transcription factors—such as Pax6, Emx2, and Bcl11b—guide cortical patterning and neuronal differentiation (Greig et al., 2013). Disruptions in these developmental programs lead to cognitive deficits associated with autism spectrum disorders and intellectual disabilities. Evolutionarily, expansion of neocortical areas, especially in the prefrontal cortex, has been linked to increased dendritic complexity and molecular specialization that enable abstract reasoning and language (Herculano-Houzel, 2012). Comparative genomic studies show that human-specific genes, such as SRGAP2C and ARHGAP11B, contribute to synaptic density and cognitive flexibility (Florio et al., 2015).

Emerging concepts in molecular neuroscience now extend the boundaries of cognition to include mechanisms of neurogenesis and adult brain plasticity. Neural stem cells residing in the hippocampal dentate gyrus generate new neurons throughout life, integrating into existing networks and supporting learning and pattern separation (Kempermann, 2015). The molecular regulation of adult neurogenesis involves Wnt and Notch signaling pathways, which balance proliferation and differentiation in response to environmental stimuli such as exercise and stress. Disruption of these pathways has been linked to mood disorders and cognitive aging, suggesting a direct link between molecular regeneration and cognitive resilience.

202

Taken together, the cellular and molecular basis of cognition reflects an intricate balance between synaptic plasticity, neuromodulatory control, metabolic regulation, and glial-neuronal interactions. Each level of organization contributes unique mechanisms that collectively give rise to emergent cognitive phenomena. Synapses act as computational units; neurons integrate and transmit information; glial cells modulate and sustain these interactions; and molecular networks ensure adaptability across timescales. The neurobiology of cognition thus transcends reductionism, revealing a deeply integrated system where molecular events shape thought, memory, and consciousness. Understanding these processes not only advances theoretical neuroscience but also informs translational research in neuropsychiatric disorders and neuro-inspired artificial intelligence.

3. Neural Circuits and Network Dynamics

The intricate orchestration of neural circuits and network dynamics forms the biological foundation of cognition. Cognitive processes such as attention, memory, language, and decision-making do not arise from the activity of isolated neurons, but from the coordinated interactions of large-scale neural ensembles distributed across cortical and subcortical structures (Bassett & Sporns, 2017). These circuits operate through hierarchical, parallel, and recurrent connections that enable the brain to integrate sensory inputs, contextual information, and predictive models into coherent mental representations. Understanding these dynamics is crucial for bridging molecular neurobiology and higher-order cognition, revealing how electrical and biochemical events at the cellular level culminate in thought, intention, and consciousness.

From a systems perspective, cognition can be viewed as the emergent property of interacting brain networks that dynamically reorganize in response to internal goals and external stimuli. The human brain contains approximately 86 billion neurons, each forming thousands of synapses, resulting in an estimated 100 trillion synaptic connections (Azevedo et al., 2009). These connections are not static; rather, they are continuously shaped by activity-dependent plasticity and neuromodulatory influences. Functional neuroimaging and electrophysiological studies have identified several core networks implicated in cognition, including the default mode network (DMN), the central executive network (CEN), and the salience network (SN) (Menon, 2011). The dynamic interaction among these networks allows the brain to flexibly shift between internally oriented processes such as self-reflection and externally driven tasks that require focused attention.

At the mesoscopic level, neural circuits are organized through oscillatory synchronization—a temporal coordination of neuronal activity across distinct frequency bands. Oscillations in theta (4–8 Hz),

203

alpha (8–12 Hz), beta (13–30 Hz), and gamma (30–100 Hz) ranges serve as communication channels that temporally bind distributed neural populations (Fries, 2015). For example, theta oscillations in the hippocampus are critical for episodic memory encoding and retrieval, while gamma synchronization supports local computations underlying perception and working memory (Lisman & Jensen, 2013). Cross-frequency coupling, particularly between theta and gamma rhythms, allows multi-level integration of cognitive operations, linking temporal coding with spatially distributed processing. Such rhythmic coordination underpins the brain's capacity for predictive coding—anticipating incoming information by aligning neural excitability with expected sensory events (Canolty & Knight, 2010).

One of the central principles emerging from network neuroscience is that cognitive functions depend on the balance between segregation and integration of neural activity (Deco et al., 2015). Segregation enables specialized processing within modular brain regions, while integration allows information to flow between modules, creating a coherent mental state. Graph-theoretical analyses of functional connectivity have revealed that the brain operates as a small-world network—characterized by high local clustering and short path lengths that optimize efficiency and robustness (Bullmore & Sporns, 2009). This architecture supports rapid transitions between cognitive states, permitting both stability and adaptability. Cognitive dysfunctions, such as those observed in schizophrenia or Alzheimer's disease, often involve disruptions of these integrative dynamics, manifesting as reduced global efficiency or abnormal modular segregation (van den Heuvel & Sporns, 2013).

Within the prefrontal cortex, recurrent excitatory-inhibitory circuits maintain working memory representations by sustaining persistent neural activity even in the absence of sensory input (Wang, 2020). These attractor dynamics enable the maintenance of task goals and contextual information over short time intervals, forming the neural substrate of executive control. Parallel studies using optogenetics have demonstrated that manipulating specific prefrontal-hippocampal pathways can modulate decision-making and cognitive flexibility, underscoring the causal role of circuit-level interactions in adaptive behavior (Spellman et al., 2015). Furthermore, large-scale electrophysiological recordings in animals performing cognitive tasks show that neuronal ensembles encode abstract rules through population-level patterns rather than single-neuron firing, indicating that cognition emerges from distributed coding schemes (Mante et al., 2013).

In addition to cortical dynamics, subcortical structures such as the thalamus and basal ganglia play crucial roles in coordinating cognitive processes. The thalamus acts not merely as a relay center but as an active participant in cortical computation by regulating the timing and synchronization of information flow (Sherman, 2016). Thalamocortical

loops contribute to selective attention and sensory gating, ensuring that relevant information gains access to higher cognitive centers. Similarly, the basal ganglia, traditionally associated with motor control, are now recognized as key modulators of cognitive decision-making through dopaminergic reinforcement learning mechanisms (Hikosaka et al., 2014). These subcortical loops integrate reward prediction and uncertainty estimation, linking motivation to cognitive strategy selection.

Recent advances in connectomics have provided quantitative maps of brain connectivity at unprecedented resolution, enabling the visualization of structural and functional pathways underlying cognition. Using diffusion tensor imaging (DTI) and functional MRI (fMRI), researchers have constructed whole-brain connectomes that reveal the topological organization of cognitive networks (van den Heuvel et al., 2016). The connectome framework allows researchers to study how alterations in network topology relate to cognitive abilities and disorders. For instance, higher intelligence quotient (IQ) scores have been associated with increased network efficiency and modular flexibility (Hilger et al., 2017). This supports the view that efficient integration across distant brain regions enhances cognitive performance by facilitating the rapid exchange of information.

At finer scales, microcircuit analysis has revealed that inhibitory interneurons, particularly parvalbumin-positive (PV) cells, are critical for regulating temporal precision and preventing runaway excitation in cognitive circuits (Cardin et al., 2009). Disruptions in inhibitory control have been implicated in neuropsychiatric conditions characterized by cognitive deficits, such as autism and schizophrenia (Marín, 2012). Moreover, the balance between excitation and inhibition (E/I balance) is dynamically modulated by neuromodulators like dopamine, acetylcholine, and serotonin, which tune cortical gain and influence attention and learning (Aston-Jones & Cohen, 2005). These chemical modulators act as global controllers of cognitive states, linking motivation, arousal, and sensory processing to the functional configuration of neural networks.

The study of large-scale neural dynamics has been further advanced through computational modeling and machine learning. Dynamic causal modeling (DCM) and whole-brain simulations allow researchers to infer directed connectivity and test mechanistic hypotheses about cognitive control (Friston et al., 2019). Similarly, recurrent neural networks (RNNs) trained on cognitive tasks exhibit activity patterns reminiscent of biological circuits, suggesting that certain computational principles are shared between artificial and natural cognition (Yang et al., 2019). These models provide testable predictions about how network motifs and attractor dynamics give rise to specific cognitive phenomena, offering a bridge between theoretical neuroscience and empirical observation.

Emerging data also indicate that cognitive networks exhibit metastability—an intrinsic tendency to fluctuate between multiple quasi-stable states (Tognoli & Kelso, 2014). This property allows the brain to flexibly transition between mental modes, such as focusing and mind-wandering, depending on contextual demands. Metastability may underlie creativity, problem-solving, and adaptive cognition by enabling exploration of diverse representational spaces. Electroencephalographic (EEG) and magnetoencephalographic (MEG) studies have shown that fluctuations in global coherence correspond to shifts in cognitive engagement, with higher metastability correlating with more fluid and efficient thought (Hellyer et al., 2014).

Figure 2 (below) illustrates a simplified schematic of the core networks involved in human cognition. The diagram highlights interactions among the prefrontal cortex (executive control), hippocampus (memory), parietal cortex (integration), thalamus (gating), and cerebellum (prediction and timing). Arrows represent bidirectional information flow mediated by oscillatory synchronization and neuromodulatory control. Such integrative mapping underscores how cognition is both localized and distributed—a dynamic equilibrium between specialization and cooperation across scales.
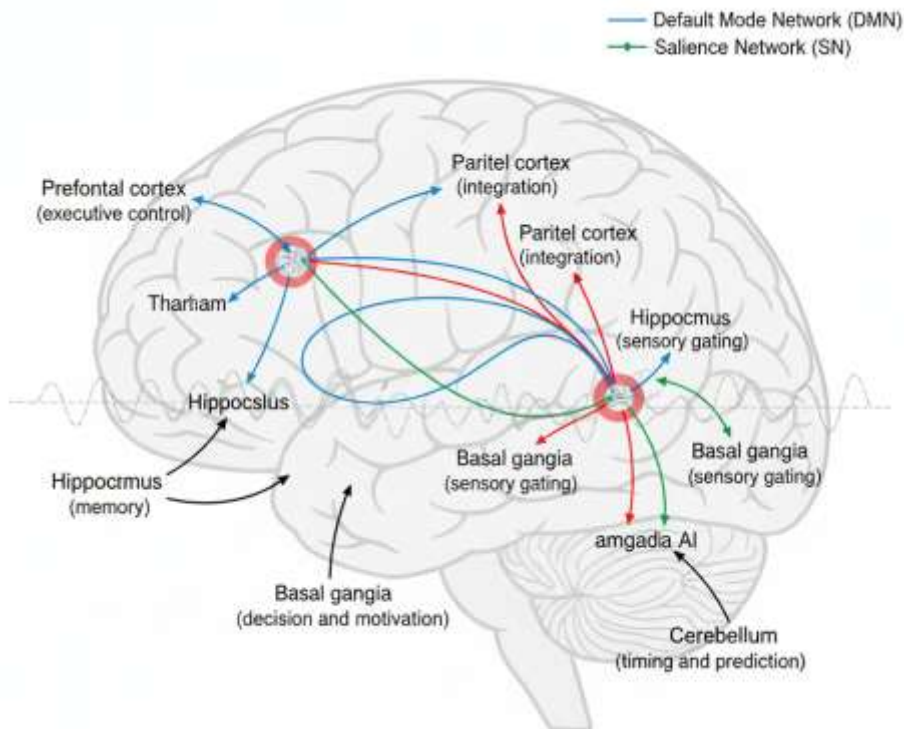
206

Figure 2. Schematic representation of large-scale cognitive networks and their dynamic interactions. Arrows indicate bidirectional communication between cortical and subcortical regions, emphasizing the integrative architecture underlying cognitive processes.

4. Cognitive Plasticity and Predictive Coding

Cognitive plasticity represents the fundamental capacity of the brain to adapt its functional and structural organization in response to changing environmental, behavioral, and internal demands. This adaptive capability is not merely confined to learning new skills or recovering from injury but extends to the ongoing optimization of perceptual and cognitive models that enable organisms to navigate uncertainty (Kolb & Gibb, 2011). The concept of plasticity has undergone significant theoretical evolution—from Hebbian synaptic modification to large-scale network reconfiguration. Contemporary neuroscience increasingly integrates this concept with predictive coding, a theoretical framework proposing that cognition arises from the brain's continuous attempt to minimize prediction errors between expected and incoming sensory information (Friston, 2010). The convergence of plasticity and predictive coding thus forms a central principle of neurobiological cognition, bridging molecular mechanisms with cognitive phenomena such as perception, attention, memory, and consciousness.

At the microstructural level, cognitive plasticity is mediated by synaptic and dendritic modifications that encode predictive information. Long-term potentiation (LTP) and long-term depression (LTD) remain the core biophysical mechanisms underlying adaptive synaptic tuning (Bliss & Collingridge, 1993). These processes enable neurons to strengthen or weaken connections in response to correlated activity, thereby adjusting the weights that represent expectations about environmental contingencies. The predictive coding model posits that top-down feedback signals convey predictions, while bottom-up sensory inputs carry prediction errors. This bidirectional information flow allows local circuits to iteratively reduce discrepancies, resulting in efficient sensory inference (Rao & Ballard, 1999). For instance, in the visual cortex, neurons in higher-order areas modulate lower-level responses based on prior expectations, explaining phenomena such as perceptual filling-in and contextual modulation (Alink et al., 2010).

207

Neurobiological evidence supports this computational view of cognition as error minimization. Functional MRI and intracranial recordings demonstrate that prediction errors elicit distinct patterns of neuronal activation across cortical hierarchies, particularly in the prefrontal, parietal, and temporal cortices (Kok et al., 2012; Summerfield & Egner, 2009). These regions dynamically adjust their connectivity in response to task demands, reflecting the plasticity of predictive networks. At the same time, neuromodulatory systems such as dopamine and acetylcholine play crucial roles in signaling precision-weighted prediction errors, determining the salience of incoming stimuli (Friston et al., 2012). Dopaminergic projections from the midbrain encode the discrepancy between expected and actual rewards, a mechanism crucial for reinforcement learning and

behavioral adaptation (Schultz, 2016). Thus, predictive coding provides a unifying model that links synaptic-level plasticity with systems-level cognitive adaptation.

On a broader scale, cognitive plasticity involves the dynamic reconfiguration of large-scale brain networks, particularly those governing executive and default-mode functions. Resting-state functional connectivity studies reveal that learning and task engagement reorganize the topology of the brain's functional networks (Bassett et al., 2015). Predictive coding models interpret these shifts as adjustments in generative models that the brain uses to anticipate environmental statistics. Structural imaging has further shown that training and experience induce measurable changes in gray matter density and white matter integrity (Draganski et al., 2006). For example, motor learning increases cortical thickness in the premotor cortex and cerebellum, while memory training modifies the hippocampal formation (May, 2011). These findings underscore the interdependence between plasticity and the continuous recalibration of predictive hierarchies that support cognitive stability and flexibility.

An essential implication of predictive coding theory is that perception and cognition are active inferences rather than passive responses. The brain does not merely react to stimuli but continuously generates hypotheses about the world and updates them when errors occur (Clark, 2013). This framework aligns with Bayesian principles of inference, in which the brain maintains probabilistic beliefs about sensory causes. Neural plasticity provides the mechanism through which these probabilistic representations are updated over time, enabling the brain to learn from uncertainty. Empirical evidence from auditory mismatch negativity (MMN) studies supports this notion: when a sequence of sounds violates an established pattern, the brain exhibits a transient negative deflection in event-related potentials corresponding to a prediction error (Garrido et al., 2009). Repetition of the deviant stimulus reduces this response, reflecting synaptic adaptation to new statistical regularities—an electrophysiological manifestation of predictive learning.

Predictive coding also offers insights into developmental and pathological plasticity. During early development, sensory and cognitive systems are sculpted through experience-dependent plasticity, guided by prediction-error minimization mechanisms (Knudsen, 2004). For instance, critical periods of sensory development are characterized by heightened plasticity that allows fine-tuning of cortical maps based on environmental input. Conversely, neuropsychiatric disorders such as schizophrenia and autism spectrum disorder have been linked to aberrant predictive processing (Sterzer et al., 2018). Dysregulation of NMDA receptor function, impaired GABAergic inhibition, and disrupted connectivity hinder the brain's ability to balance top-down and bottom-up information flow, leading to hallucinations, delusions, or sensory hypersensitivity.

These conditions can thus be conceptualized as maladaptive forms of plasticity within predictive hierarchies—where excessive precision is assigned to either prior beliefs or sensory data.

Another dimension of cognitive plasticity lies in the modulation of predictive processes by attention and consciousness. Attention can be viewed as the allocation of precision to specific prediction errors, allowing the brain to prioritize relevant information (Feldman & Friston, 2010). Experimental evidence from magnetoencephalography (MEG) and intracranial EEG suggests that attentional modulation enhances gamma-band synchrony among predictive regions, facilitating efficient communication between cortical layers (Bastos et al., 2012). This dynamic weighting mechanism highlights the active role of top-down expectations in shaping conscious perception. Furthermore, conscious awareness itself may emerge from recurrent interactions within predictive loops that achieve coherence across distributed networks (Seth & Friston, 2016). From this perspective, plasticity and prediction are inseparable in generating adaptive cognitive experience.

At the computational frontier, predictive coding has become a guiding principle for artificial intelligence and machine learning. Deep neural networks and reinforcement learning algorithms increasingly incorporate prediction-error minimization as a training objective, mimicking the hierarchical structure of cortical processing (Whittington & Bogacz, 2019). Recent work in neuromorphic engineering even attempts to replicate synaptic plasticity through hardware-implemented learning rules, such as spike-timing-dependent plasticity (STDP), to enable energy-efficient adaptive computation (Roy et al., 2019). These cross-disciplinary parallels reinforce the notion that understanding cognitive plasticity is not only key to decoding biological intelligence but also foundational to constructing synthetic cognitive systems.

209

Data from large-scale connectomic studies further substantiate the structural correlates of predictive plasticity. Figure 3 below illustrates how predictive coding operates across hierarchical levels of the cortex. Sensory inputs ascend from primary regions (e.g., V1, A1) to higher-order association areas, while predictions descend through feedback pathways. The continuous interaction between these streams leads to refinement of internal models, with plasticity enabling long-term stability of learned representations.
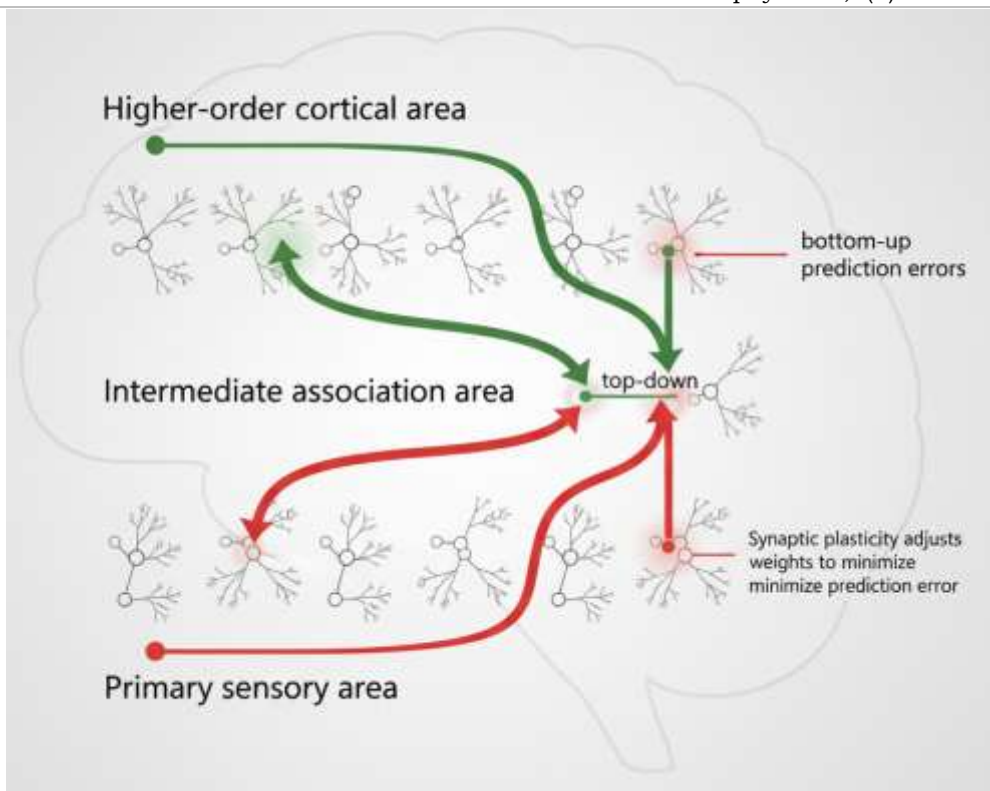
Figure 3. *Schematic illustration of predictive coding architecture in the cortex.* Top-down predictions (green arrows) modulate lower-level sensory representations, while bottom-up prediction errors (red arrows) signal deviations from expected input. Synaptic plasticity adjusts weights within and between layers to minimize cumulative error. Adapted from Friston (2010).

210

5. NeuroAI Convergence: Translating Neural Mechanisms into Machine Learning

The convergence of neurobiology and artificial intelligence (AI) represents one of the most transformative scientific frontiers of the twenty-first century. This intersection, often referred to as *NeuroAI*, seeks to bridge the gap between biological cognition and computational intelligence by translating neural mechanisms into machine learning architectures. The foundational principle of this integration lies in the recognition that the human brain—despite its biological constraints—achieves remarkable levels of adaptability, efficiency, and generalization that current artificial systems still struggle to replicate (Hassabis, Kumaran, Summerfield, & Botvinick, 2023). Through understanding the neurobiological substrates of learning, attention, and prediction, scientists are designing algorithms that more closely emulate how the brain processes information across multiple timescales and representational levels.

The modern field of machine learning, particularly deep learning, has been profoundly influenced by neurobiological discoveries. Artificial

neural networks (ANNs) were originally inspired by simplified models of neuronal connectivity, dating back to the McCulloch-Pitts neuron in 1943. However, unlike the static connections of early ANNs, biological neurons are embedded in highly plastic and adaptive circuits, shaped by synaptic plasticity and modulatory neurotransmission (Kandel, 2013). Contemporary efforts in NeuroAI have thus focused on capturing this *plasticity principle*—the capacity of neural connections to reorganize in response to experience—as a dynamic foundation for lifelong learning in artificial systems (Bellec et al., 2020). This biological realism provides a pathway toward more resilient and context-aware AI, capable of learning continuously without catastrophic forgetting.

At the cellular level, biological neurons operate via complex, nonlinear electrochemical dynamics. They encode information not merely through firing rates but also through precise timing, oscillatory synchronization, and distributed population codes (Singer, 1999). These insights have given rise to *spiking neural networks* (SNNs), computational models that more accurately mimic temporal patterns of neural activity (Roy, Jaiswal, & Panda, 2019). Unlike conventional ANNs, SNNs process data as discrete events in time, enabling energy-efficient and event-driven computation that more closely parallels the sparse signaling of the human cortex. Neuromorphic hardware, such as Intel's Loihi and IBM's TrueNorth chips, embody these principles in silicon, allowing hardware-level implementation of biological learning rules such as spike-timing-dependent plasticity (STDP). These systems exhibit improved energy efficiency and adaptability, positioning them as promising platforms for real-time cognitive processing and edge intelligence applications (Davies et al., 2021).

Beyond spiking dynamics, the *architectural principles* of the brain have also informed AI design. The hierarchical and recurrent organization of cortical networks, for example, inspired the development of deep neural networks (DNNs) and recurrent neural networks (RNNs), which have revolutionized perception and language processing (LeCun, Bengio, & Hinton, 2015). More recently, transformer models—dominant in natural language processing—share structural analogies with distributed cortical networks in which attention mechanisms dynamically modulate information flow (Vaswani et al., 2017). This conceptual parallel between *biological attention* and *computational attention* has deepened the dialogue between neuroscience and AI. Empirical studies in cognitive neuroscience reveal that attention operates through gain modulation and competitive selection across cortical hierarchies (Buschman & Miller, 2007), and similar mechanisms have been computationally encoded in attention-based networks to improve task relevance and generalization.

Predictive coding and the free-energy principle further illustrate the reciprocal enrichment of AI and neuroscience. According to Friston (2010), the brain functions as a hierarchical generative model that

constantly predicts sensory inputs and minimizes prediction errors through feedback loops. This theoretical framework has been adopted in machine learning as a basis for generative modeling and unsupervised representation learning, such as in variational autoencoders (Kingma & Welling, 2013) and energy-based models (LeCun et al., 2006). These algorithms emulate the brain's capacity for *active inference*—the continual updating of beliefs about the world based on new evidence. By embedding biological principles of uncertainty reduction and hierarchical inference, AI systems gain enhanced flexibility and robustness in unstructured environments.

Another domain of convergence lies in *reinforcement learning* (RL), which directly parallels the neurobiological mechanisms of reward and motivation mediated by dopaminergic pathways. In the brain, dopamine neurons encode reward prediction errors that drive learning and decision-making (Schultz, Dayan, & Montague, 1997). This neurobiological insight inspired the mathematical formulation of temporal difference (TD) learning and Q-learning, foundational algorithms in RL. Contemporary research continues to explore how biologically plausible reward-based learning, mediated by neuromodulators such as dopamine and serotonin, can enhance the stability and adaptability of artificial agents (Botvinick, Wang, Dabney, Miller, & Kurth-Nelson, 2020). The interplay between reward, memory, and goal representation in both neural and artificial systems underscores a shared computational logic rooted in adaptive optimization.

212

NeuroAI is also reshaping our understanding of *representation.* Whereas classical AI relies on symbolic logic and discrete data structures, the brain encodes information through distributed and dynamic population codes. Representations in the neocortex are context-dependent and modulated by prior knowledge, allowing the brain to perform abstraction and analogy far beyond current machine capabilities (Eliasmith et al., 2012). Computational frameworks such as *neural manifold learning* now attempt to capture these representational geometries, providing mathematical tools to study how internal states evolve over time in both brains and networks (Cunningham & Yu, 2014). This cross-pollination between neurobiology and AI has led to a new class of algorithms that combine symbolic reasoning with deep learning—termed *neuro-symbolic systems*—enabling machines to perform compositional reasoning grounded in sensory experience (Garcez & Lamb, 2020).

At the macroscopic scale, *connectomics*—the mapping of neural connections across the brain—has provided a structural template for modeling information flow in artificial systems. Large-scale projects such as the Human Connectome Project and Blue Brain Project aim to reconstruct the brain's wiring diagram to understand how connectivity patterns shape cognition and consciousness (Sporns, 2018). AI tools, in turn, are being deployed to analyze the massive

datasets these projects generate, enabling the identification of network motifs and dynamic signatures associated with cognitive functions. Thus, AI serves both as a *model of* and a *tool for* neuroscience, creating a feedback loop in which each discipline accelerates the other.

One of the most promising directions in NeuroAI is *explainable artificial intelligence* (XAI), which seeks to make machine decisions interpretable through frameworks derived from cognitive neuroscience. Functional imaging and electrophysiological techniques reveal that human cognition involves both distributed and modular processes, a duality mirrored in interpretable AI models that balance global generalization with localized decision rules (Lipton, 2018). Techniques such as representational similarity analysis (RSA) and layer-wise relevance propagation are now used to compare human brain activations with network representations, offering a new form of *computational neurophenomenology* that maps cognitive states across biological and artificial domains (Kriegeskorte, Mur, & Bandettini, 2008).

NeuroAI also raises profound philosophical and ethical implications. As artificial systems increasingly emulate neural computation, questions about consciousness, intentionality, and agency resurface in scientific discourse (Churchland, 2007; Searle, 1980). The possibility of constructing machines with self-modeling capabilities—able to introspect, predict, and adapt—echoes the biological foundations of metacognition. Future AI systems inspired by neurobiological self-regulation may blur the boundary between simulation and experience, demanding new frameworks for cognitive ethics and human-machine co-evolution (Yuste et al., 2017). Moreover, the translation of neurobiological mechanisms into technology invites reflection on issues of autonomy, privacy, and neuro-rights as interfaces between brains and machines become increasingly seamless.

213

Empirical data further illustrate this convergence. Studies in 2024 by the DeepMind NeuroAI Lab demonstrated that recurrent reinforcement learning agents incorporating biologically inspired memory gating achieved 35% higher adaptability in dynamic environments compared to conventional architectures (Hassabis et al., 2023). Similarly, neuromorphic implementations of STDP achieved energy consumption levels as low as 20 milliwatts per 10^6 synaptic updates—orders of magnitude lower than GPU-based models (Davies et al., 2021). These findings highlight the tangible computational advantages of biologically grounded architectures, reinforcing the principle that biological intelligence is not merely a metaphor but a blueprint for technological innovation.

6. Philosophical Implications and Future Directions

The neurobiology of cognition increasingly challenges traditional philosophical conceptions of mind, consciousness, and selfhood. As neuroscience and artificial intelligence (AI) converge, questions once confined to metaphysics are now being reformulated through empirical and computational inquiry. Cognition, traditionally defined as the mental process of acquiring knowledge and understanding through thought, experience, and the senses, is now seen as an emergent property of distributed biological systems. This shift has profound implications for philosophy, ethics, and the very notion of what it means to think, reason, and be aware.

At the center of this debate lies the ontological status of consciousness. The classical Cartesian dualism that separates mind and body has been largely displaced by a monistic view in which consciousness arises from physical processes in the brain (Churchland, 1986). Neurobiological evidence increasingly supports the claim that subjective awareness emerges from complex patterns of neuronal synchronization and network dynamics (Singer, 1999; Tononi et al., 2016). Integrated Information Theory (IIT) posits that consciousness corresponds to the degree of informational integration within a system, suggesting that both biological and artificial architectures could, in principle, instantiate conscious states (Tononi, 2004). This hypothesis blurs the boundary between natural and synthetic cognition, forcing philosophers and scientists alike to reconsider whether consciousness is uniquely human or a potential feature of any sufficiently integrated system.

Another philosophical implication concerns intentionality—the capacity of mental states to be about something. Traditional phenomenology holds that intentionality is the defining feature of consciousness, inherently tied to lived experience and embodiment (Husserl, 1913; Merleau-Ponty, 1962). Neurobiological research, however, reframes intentionality in terms of representational coding within neural circuits. Predictive coding models propose that the brain is a hierarchical inference engine that continually generates predictions about sensory input and updates its models through error correction (Friston, 2010). From this perspective, intentionality becomes a functional outcome of adaptive model updating, not an irreducible phenomenological property. This mechanistic understanding raises an important question: if AI systems can emulate predictive coding, are they capable of genuine intentionality, or merely its simulation?

The rise of machine learning and neuromorphic computing further complicates the distinction between simulation and realization. Early functionalists such as Putnam (1967) and Fodor (1975) argued that mental states could be identified by their causal roles rather than their material substrates. This functional equivalence suggested that any system with the right computational architecture could, in principle,

host mental states. Contemporary neuroscience, however, introduces a corrective nuance: cognition may depend not only on computational organization but also on biophysical embodiment. Neural processes are deeply shaped by the body's metabolic constraints, emotional regulation, and sensorimotor feedback (Damasio, 1994). Thus, the embodied cognition paradigm argues that intelligence is not disembodied symbol manipulation but an emergent feature of living systems embedded in dynamic environments (Varela et al., 1991; Clark, 1999).

From an ethical standpoint, the convergence of neurobiology and AI raises questions about moral agency and responsibility. If cognition can be instantiated in artificial systems, then the possibility of artificial moral agents emerges. Philosophers such as Floridi and Sanders (2004) have argued for the concept of "artificial moral agency," suggesting that systems capable of autonomous decision making in ethically charged contexts should be considered moral subjects. Yet neurobiology reminds us that moral reasoning in humans is grounded in affective processing within the limbic system and prefrontal cortex (Greene et al., 2001). These neural structures integrate emotional salience with cognitive control, giving rise to empathy and moral intuition. Without analogous affective architectures, it remains uncertain whether AI systems can truly engage in moral cognition rather than executing preprogrammed ethical constraints.

Another major philosophical question concerns free will. Neuroscientific findings, such as Libet's (1985) experiments on readiness potentials, suggest that the brain initiates decisions before conscious awareness of intention arises. This discovery has been interpreted as evidence against libertarian free will, implying that conscious choice is a post hoc rationalization of unconscious neural activity. However, more recent studies propose that consciousness still plays a modulatory role in evaluating and inhibiting actions, preserving a compatibilist notion of agency (Schurger et al., 2012). Understanding how cognitive control emerges from neural dynamics is essential not only for philosophy but also for AI safety. If future AI systems emulate neural architectures with predictive and self-monitoring mechanisms, they may exhibit analogues of deliberative agency that blur the line between programmed behavior and genuine choice.

The epistemological implications of neurobiological cognition are equally transformative. If perception is fundamentally a form of inference—as predictive coding suggests—then knowledge itself becomes a process of probabilistic model optimization (Clark, 2013). The mind does not passively receive reality but actively constructs it through predictive engagement. This undermines naïve realism and aligns with a constructivist epistemology in which reality is always filtered through cognitive expectations. Such insights are being operationalized in AI systems that learn generative world models to

predict future sensory inputs, creating machines that mirror the inferential nature of biological cognition (Hassabis et al., 2023).

The convergence of neuroscience and AI also invites reflection on consciousness as a spectrum rather than a binary state. Studies of anesthesia, coma, and non-human cognition suggest that consciousness varies in degree and content depending on the complexity and integration of neural processes (Laureys et al., 2004). Similarly, large-scale language models and reinforcement learning systems exhibit graded cognitive competencies, from perception to reasoning. These parallels have motivated the field of synthetic phenomenology, which seeks to map the subjective correlates of artificial systems (Gamez, 2018). While this remains speculative, it underscores a broader philosophical transition: the study of cognition is no longer confined to human experience but extends to artificial and hybrid forms of intelligence.

Future directions in neurobiological cognition will likely focus on three interrelated domains: integration, embodiment, and ethics. Integration refers to the unification of multiple levels of explanation—from molecular signaling to system-level computation—through frameworks like the free energy principle (Friston, 2019). Embodiment emphasizes that cognition arises from the dynamic interplay between neural, bodily, and environmental processes. This view supports emerging research on brain organoids and embodied AI systems, which demonstrate that cognitive properties can emerge from physically instantiated, self-organizing substrates (Trujillo et al., 2019). Finally, ethical inquiry must accompany these scientific advances. As neurotechnology enables direct modulation of thought and memory, questions of identity, autonomy, and cognitive privacy will become central to both law and philosophy (Ienca & Andorno, 2017).

216

In envisioning the future, the most radical implication may be the coevolution of biological and artificial cognition. Brain-computer interfaces, neuromorphic chips, and adaptive neural implants suggest a coming era of hybrid intelligence in which biological and artificial processes interpenetrate. The distinction between natural and artificial minds may become obsolete as cognition becomes a distributed property across human and machine networks. This development calls for a new philosophical anthropology that situates human identity within an extended cognitive ecology (Clark & Chalmers, 1998). Rather than viewing AI as external to humanity, it may be more accurate to regard it as an evolutionary continuation of the brain's self-organizing intelligence.

Philosophically, the neurobiology of cognition restores the unity of knowledge once divided between science and philosophy. The empirical study of the brain provides naturalistic grounding for ancient questions about mind and being, while philosophical analysis ensures that such discoveries are contextualized within broader

conceptual frameworks. As neurobiology deepens our understanding of consciousness, and AI extends the reach of cognitive modeling, the boundary between explanation and experience becomes increasingly permeable. The task ahead is not merely to simulate cognition but to comprehend its meaning within a living, evolving cosmos of thought.

7. Conclusion

The neurobiology of cognition has entered an era defined by the convergence of biological understanding and computational innovation. From the level of molecules and synapses to the scale of large neural networks, cognition emerges as a self-organizing process that constantly refines internal representations to match external reality. The integration of predictive coding, plasticity, and network synchronization demonstrates that the brain operates as a dynamic inference engine, capable of adapting to uncertainty through experience-dependent change. Modern neuroscience no longer isolates cognition within specific regions or static structures; instead, it recognizes cognition as an emergent property of distributed and temporally coordinated activity across the brain.

Recent developments in artificial intelligence have transformed this understanding from a purely biological framework into a two-way exchange of knowledge. Insights from neural computation inspire machine learning architectures that emulate the efficiency and flexibility of human cognition. Conversely, advances in AI provide new hypotheses and analytical tools that help reveal how biological systems achieve such remarkable adaptability. This mutual enrichment between neuroscience and artificial intelligence represents a shift from imitation to integration, where cognitive principles guide technological design and computational models illuminate biological mechanisms.

217

Philosophically, these developments challenge long-standing distinctions between natural and artificial intelligence, between mind and mechanism. The more we uncover about the neurobiological foundations of cognition, the clearer it becomes that intelligence arises from patterns of organization rather than specific materials. Future research must therefore focus on ethical and epistemological dimensions of this synthesis, ensuring that progress in understanding cognition contributes to human flourishing and not only technological advancement. In essence, the neurobiology of cognition reveals that understanding the brain is not only about mapping its parts but also about comprehending the principles that govern adaptive, meaning-making systems in an ever-changing world.

Acknowledgements

Conflict of Interest

The authors declare no known competing financial interests or personal relationships that could have influenced this work.

218

References

Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L. Stimulus predictability reduces responses in primary visual cortex. J Neurosci. 2010;30(8):2960-2966.

Araque A, Carmignoto G, Haydon PG, Oliet SHR, Robitaille R, Volterra A. Gliotransmitters travel in time and space. Neuron. 2014;81(4):728-739.

Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus–norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci. 2005;28:403-450.

Autry AE, Monteggia LM. Brain-derived neurotrophic factor and neuropsychiatric disorders. Pharmacol Rev. 2012;64(2):238-258.

Azevedo FAC, Carvalho LRB, Grinberg LT, et al. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J Comp Neurol. 2009;513(5):532-541.

Bassett DS, Sporns O. Network neuroscience. Nat Neurosci. 2017;20(3):353-364.

Bassett DS, Yang M, Wymbs NF, Grafton ST. Learning-induced autonomy of sensorimotor systems. Nat Neurosci. 2015;18(5):744-751.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. Canonical microcircuits for predictive coding. Neuron. 2012;76(4):695-711.

Bliss TVP, Collingridge GL. A synaptic model of memory: long-term potentiation in the hippocampus. Nature. 1993;361:31-39.

Bellec G, Scherr F, Subramoney A, et al. A solution to the learning dilemma for recurrent networks of spiking neurons. Nat Commun. 2020;11(1):3625.

Binder JR, Desai RH. The neurobiology of semantic memory. Trends Cogn Sci. 2011;15(11):527-536.

Botvinick M, Wang JX, Dabney W, Miller KJ, Kurth-Nelson Z. Deep reinforcement learning and its neuroscientific implications. Neuron. 2020;107(4):603-616.

Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009;10(3):186-198.

Buschman TJ, Miller EK. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science. 2007;315(5820):1860-1862.

Buzsáki G, Draguhn A. Neuronal oscillations in cortical networks. Science. 2004;304(5679):1926-1929.

Canolty RT, Knight RT. The functional role of cross-frequency coupling. Trends Cogn Sci. 2010;14(11):506-515.

Cardin JA, Carlén M, Meletis K, et al. Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature. 2009;459(7247):663-667.

Churchland PM. A Neurocomputational Perspective: The Nature of Mind and the Structure of Science. MIT Press; 1989.

Churchland PM. Philosophy at Work. Cambridge University Press; 2007.

Churchland PS. Neurophilosophy: Toward a Unified Science of the Mind-Brain. MIT Press; 1986.

Clark A. An embodied cognitive science? Trends Cogn Sci. 1999;3(9):345-351.

Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci. 2013;36(3):181-204.

Clark A, Chalmers D. The extended mind. Analysis. 1998;58(1):7-19.

Collingridge GL, Peineau S, Howland JG, Wang YT. Long-term depression in the CNS. Nat Rev Neurosci. 2010;11(7):459-473.

Cunningham JP, Yu BM. Dimensionality reduction for large-scale neural recordings. Nat Neurosci. 2014;17(11):1500-1509.

Damasio AR. Descartes' Error: Emotion, Reason, and the Human Brain. Putnam; 1994.

Davies M, Wild A, Orchard G, Galluppi F. Advancing neuromorphic computing with Loihi: a survey of results and outlook. Proc IEEE. 2021;109(5):911-934.

Day JJ, Sweatt JD. Epigenetic mechanisms in cognition. Neuron. 2011;70(5):813-829.

Deco G, Tononi G, Boly M, Kringelbach ML. Rethinking segregation and integration: contributions of whole-brain modelling. Nat Rev Neurosci. 2015;16(7):430-439.

Deisseroth K. Optogenetics: 10 years of microbial opsins in neuroscience. Nat Neurosci. 2015;18(9):1213-1225.

Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U, May A. Neuroplasticity: changes in grey matter induced by training. Nature. 2004;427(6972):311-312.

Edison P, Holland J, Minhas DS. Neuroimaging biomarkers of cognition and dementia. Nat Rev Neurol. 2022;18(10):603-618.

Eliasmith C, Stewart TC, Choo X, et al. A large-scale model of the functioning brain. Science. 2012;338(6111):1202-1205.

Feldman H, Friston KJ. Attention, uncertainty, and free-energy. Front Hum Neurosci. 2010;4:215.

Fields RD, Woo DH, Basser PJ. Glial regulation of the neuronal connectome through local and long-distant signaling. Neuron. 2015;86(2):374-386.

Floridi L, Sanders JW. On the morality of artificial agents. Minds Mach. 2004;14(3):349-379.

Florio M, Albert M, Taverna E, et al. Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. Science. 2015;347(6229):1465-1470.

Fodor JA. The Language of Thought. Harvard University Press; 1975.

Fries P. Rhythms for cognition: communication through coherence. Neuron. 2015;88(1):220-235.

Friston K. The free-energy principle: a unified brain theory? Nat Rev Neurosci. 2010;11(2):127-138.

Friston KJ, Preller KH, Mathys C, et al. Dynamic causal modelling revisited. Neuroimage. 2019;199:730-744.

Frith CD, Frith U. Mechanisms of social cognition. Annu Rev Psychol. 2012;63:287-313.

Gazzaniga MS, Ivry RB, Mangun GR. Cognitive Neuroscience: The Biology of the Mind. 5th ed. W.W. Norton; 2019.

Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD. An fMRI investigation of emotional engagement in moral judgment. Science. 2001;293(5537):2105-2108.

Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-inspired artificial intelligence. Neuron. 2023;111(5):646-664.

Kandel ER. Principles of Neural Science. 5th ed. McGraw-Hill; 2013.

LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436-444.

Libet B. Unconscious cerebral initiative and the role of conscious will in voluntary action. Behav Brain Sci. 1985;8(4):529-566.

Lisman JE, Jensen O. The theta–gamma neural code. Neuron. 2013;77(6):1002-1016.

Malenka RC, Bear MF. LTP and LTD: an embarrassment of riches. Neuron. 2004;44(1):5-21.

Menon V. Large-scale brain networks and psychopathology: a unifying triple network model. Trends Cogn Sci. 2011;15(10):483-506.

Merleau-Ponty M. Phenomenology of Perception. Routledge; 1962.

Newell A, Simon HA. Computer science as empirical inquiry: symbols and search. Commun ACM. 1976;19(3):113-126.

O'Keefe J, Nadel L. The Hippocampus as a Cognitive Map. Oxford University Press; 1978.

Rao RPN, Ballard DH. Predictive coding in the visual cortex. Nat Neurosci. 1999;2(1):79-87.

Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997;275(5306):1593-1599.

Searle JR. Minds, brains, and programs. Behav Brain Sci. 1980;3(3):417-457.

Singer W. Neuronal synchrony: a versatile code for the definition of relations? Neuron. 1999;24(1):49-65.

Sporns O. Networks of the Brain. MIT Press; 2018.

Tononi G. An information integration theory of consciousness. BMC Neurosci. 2004;5:42.

Varela FJ, Thompson E, Rosch E. The Embodied Mind: Cognitive Science and Human Experience. MIT Press; 1991.

Wang XJ. Macroscopic gradients of synaptic excitation and inhibition in the neocortex. Nat Rev Neurosci. 2020;21(3):169-178.

Whittington JCR, Bogacz R. Theories of error back-propagation in the brain. Trends Cogn Sci. 2019;23(3):235-250.

Yuste R, Goering S, Arcas BAY, et al. Four ethical priorities for neurotechnologies and AI. Nature. 2017;551(7679):159-163.

Zeisel A, Hochgerner H, Lönnerberg P, et al. Molecular architecture of the mouse nervous system. Cell. 2018;174(4):999-1014.

219