

Part IV — The Biological Blueprint

Contents

Part IV: The Biological Blueprint: Evidence for UAF.	1
Chapter 28: The Human Brain: A Living Blueprint of UAF.	1
Chapter 29: Evolutionary Drivers: Skin in the Game in Biological Systems.	2
Chapter 30: The Architecture of Biological Qualia: Insights from Cognitive Science.	4
Chapter 31: Mental Illness as a Failure of Functional Fiction: A UAF Perspective.	5
Key References Cited (<i>Harvard Style, Alphabetical</i>)	6

Part IV: The Biological Blueprint: Evidence for UAF.

Chapter 28: The Human Brain: A Living Blueprint of UAF.

Having explored UAF’s theoretical framework and its power in re-framing philosophical puzzles, we now turn to ourselves. We have the only consciousness that we truly know well. We have long thought that consciousness is what makes us special. Behind our consciousness is the human brain. The living mass of cells that are deeply interconnected forming a network of communication and information processing. It is something that from the point of view of UAF is the incredibly efficient and complex **Underlying Computational System (UCS)** providing the power for the formation of the representation of ourselves, the universe and the interaction between these two. The approximation of what it is like to be a human interacting with the world.

The human brain, with its **billions of neurons and trillions of synaptic connections** (Herculano-Houzel, 2009), is currently the most impressive and beautiful form of information processing. Yet, even with its most magnificent place on the known information processing systems, it cannot possibly process every quantum fluctuation or every molecular interaction occurring within itself or its environment. *This inherent limitation is a fundamental aspect of any finite system, leading to what is sometimes called the “binding problem” in neuroscience—how disparate neural activities are integrated into a unified perception without overwhelming the system (Treisman, 1996).* The brain is beautiful, complex and powerful, but it is very small compared to the universe. This inherent complexity forces the brain to find simplifications. Most of the details of our surroundings can be ignored in our daily lives. We do not need to understand all the fusion reactions and the detailed flow of hydrogen and helium atoms in our sun to understand that we circle around it and it gives us light during the day and is hidden behind the planet during the night. We do not need to know about the molecules on the surface of strawberries to understand what it looks like and tastes like. We do not consciously experience the firing of individual neurons, the reactions to neurotransmitters and detailed flow of information in our brain. Instead, our consciousness is presented with higher-level, simplified approximations — the thoughts, perceptions, and feelings that constitute our subjective reality (Metzinger, 2009; Dennett, 1991). This is the brain’s elegant solution to avoid **Computational Paralysis** (Hofstadter, 1979).

At its core, the brain operates as a vast, interconnected **neural network**. This network is not merely a passive receiver of information; it is an active, predictive engine. Sensory processing, for instance, is not a simple bottom-up relay of data. Instead, the brain constantly generates predictions about incoming sensory input, comparing these predictions to the actual signals received. This is the essence of **predictive coding**, a widely accepted model in modern neuroscience (Friston, 2010; Clark, 2016). When there’s a mismatch — a **prediction error** — the brain updates its internal models to reduce future discrepancies. This process of **Prediction Error Minimization (PEM)** is the fundamental learning mechanism that refines the brain’s approximations of reality (Hohwy, 2013).

This predictive architecture directly gives rise to the core components of UAF:

- **The World-Model:** The brain continuously constructs and refines an internal **World-Model**—a dynamic, approximate representation of its external environment. This model is built and adjusted to match the vast stream of sensory data. The model has been getting more and more accurate to match reality every year of your life. From the moment we are born, our brains are busy mapping the spatial layout of our surroundings, the properties of objects, the behavior of other beings, and the causal relationships that govern the world (Barsalou, 2008). This World-Model is not a perfect replica, but a functionally optimized approximation, allowing us to navigate, predict, and interact with our environment efficiently. It is especially tuned to match your focus, needs and interests. *This constant refinement of the World-Model is evident in phenomena like **perceptual learning**, where repeated exposure to stimuli leads to more accurate and efficient processing (Fahle, 2005).*
- **The Internal Self-Model (ISM):** Simultaneously, the brain builds an **Internal Self-Model (ISM)** — its own user interface, a simplified approximation of its own body, reflexes, subconscious reactions, and its own behavior. This ISM integrates proprioceptive (body position), interoceptive (internal body states like hunger or heart rate), motor (action-related) signals, hormonal reactions to the surrounding, automatic reactions of the brain and even the workings of our thought processes (Damasio, 1999; Craig, 2002). It’s how the brain knows “this hand is mine” or “I am tired.” The self-model also contains a representation of itself as information processing system: we understand that we receive information through our senses, process it and integrate it with our memories and allow it to form new memories and finally use all of our knowledge of the current and the past to make decisions and react to the environment. This self-model is crucial for agency, allowing the brain to plan and execute actions coherently without getting bogged down in the microscopic details of its own neural machinery (Metzinger, 2003).
- **Qualia:** The subjective “feelings” or “phenomenal flavors” that we experience—our **Qualia**—are the brain’s “simplified truths.” They are the highly compressed, functionally essential interpretations of complex internal and external information (Seth, 2021; Chalmers, 1996). The searing pain of a burn, the vibrant hue of red, the feeling of joy—these are not raw neural firings, but the brain’s ultimate summary signals, providing **Subjective Closure** and driving **Causal Efficacy**. They are the dashboard indicators that allow the brain to immediately understand and respond to critical information, ignoring layers of detailed processing. *This compression is vital for rapid decision-making, allowing the system to prioritize and respond to salient features of its environment without being overwhelmed by raw data (Kahneman, 2011).*

The brain’s entire architecture is geared towards managing complexity and enabling coherent action. From the hierarchical processing in sensory cortices (where simple features combine into complex objects) (Marr, 1982; Felleman & Van Essen, 1991) to the intricate feedback loops between different brain regions, every aspect points to a system designed to build and refine approximations. The brain doesn’t strive for absolute truth; it strives for *useful* truth — the most efficient, actionable approximation that maximizes the likelihood of survival and propagation.

Consider the brain’s remarkable ability to fill in missing information or create coherent perceptions from ambiguous input. When we see a partially obscured object, our brain doesn’t just see fragments; it uses its World-Model to predict and “fill in” the missing parts, creating a complete, albeit approximate, perception. This “controlled hallucination,” as some neuroscientists describe it (Seth, 2021), is a direct manifestation of the brain’s predictive, approximate nature. It’s not about seeing reality as it is, but about constructing the most probable and useful reality given limited, noisy data (Hohwy, 2013).

The brain is the living truth behind the approximation described in this book and behind the idea of UAF. It is also the inspiration for complex neural networks, which are a very simplified approximation of what the brain is. As a living system, the brain has a lot of additional machinery supporting their main function. Some of it has an effect on the main signal that the neurons are transmitting, but it is mostly a minimal noise signal. *Beyond neurons, **glial cells** (astrocytes, oligodendrocytes, microglia) play crucial roles in modulating synaptic activity, providing metabolic support, and influencing neural plasticity, demonstrating that the UCS is a highly integrated, multi-component system (Fields, 2009).*

Chapter 29: Evolutionary Drivers: Skin in the Game in Biological Systems.

The abstract imperative of ‘**Skin in the Game**’ (SiG) (Taleb, 2018), which compels systems towards survival is one of the main components of biological evolution. Life on Earth, from its earliest microbial

forms to the most complex human societies, is a continuous struggle for existence, driven by the fundamental need to survive, reproduce, and pass on genetic instructions (Darwin, 1859; Dawkins, 1976). This existential pressure is the ultimate form of Skin in the Game, providing the evolutionary force that pushes for the formation of subconscious behavioral patterns and learning part of the brain that then forms the approximations of what is reality, what the brain as a whole is, and ultimately leading to the emergence of consciousness as the representation of what it is like to be such a being.

Natural selection, the engine of evolution, is a brutal and unforgiving accountant of efficiency. Organisms that are better at predicting their environment, finding resources, avoiding predators, and successfully reproducing are the ones whose genes propagate. Those that fail to do so are culled. This constant, high-stakes feedback loop creates immense **Skin in the Game** for every living being. It's not an optional game; it's the only game in town, with the ultimate stakes: existence or extinction.

This intense pressure drives the **Imperative for Coherence & Agency**. A simple bacterium needs a rudimentary form of coherence to move towards nutrients and away from toxins. A complex mammal needs a far more sophisticated level of coherence to navigate a vast territory, hunt prey, evade predators, and raise offspring (Sterelny, 2003). Every cell of the mammal has its own struggle to be useful in its micro environment. Liver cells try to balance their environment by detoxifying harmful substances, processing nutrients, and maintaining metabolic homeostasis, while muscle cells strive to generate force and facilitate movement, and neurons work to transmit signals and coordinate complex behaviors. Each cell type contributes to the overall coherence and agency of the organism, ensuring its survival and ability to thrive in its environment. *This cellular-level SiG is governed by **gene regulatory networks** that ensure cells specialize and cooperate, forming a coherent multicellular organism (Davidson, 2006).* This imperative, born from SiG, pushes for the development of more efficient information processing systems — neurons and brains — that can build better, more useful approximations of reality.

Consider the evolutionary pressure points that directly led to the formation of the primitive shared brain structures and behavioral patterns found in most mammals:

- **Resource Scarcity:** Food, water, and shelter are rarely abundant. Organisms with better **World-Models** (Chapter 9) that can accurately predict the location of resources, remember past foraging successes, and anticipate seasonal changes have a distinct survival advantage (Shettleworth, 2010). This drives the evolution of memory, spatial navigation (e.g., hippocampus in mammals), and predictive capabilities.
- **Predation:** Being eaten is the ultimate form of losing your Skin in the Game. Organisms that can quickly and accurately identify threats, predict predator behavior, and execute rapid escape responses are more likely to survive. This pushes for the evolution of sophisticated sensory processing, rapid decision-making, and the generation of urgent **Qualia** like fear or pain, which provide fast “simplified truths” that compel action (LeDoux, 1996).
- **Reproduction:** Passing on genes is the ultimate measure of evolutionary success. This involves finding mates, successfully reproducing, and often, raising offspring. These complex social and biological tasks require sophisticated **Internal Self-Models (ISM)** (Chapter 7) (e.g., understanding one's own physical state, social standing, and reproductive readiness) and refined World-Models (e.g., understanding potential mates, social hierarchies, and offspring needs) (Buss, 1994). The pleasure qualia associated with successful reproduction or social bonding are powerful drivers, reinforcing behaviors that lead to genetic propagation (Rolls, 2000).
- **Competition:** Organisms compete not just with other species, but with their own kind for mates and resources. This drives the evolution of social intelligence, deception, cooperation, and the ability to predict the behavior of rivals—all requiring increasingly complex and nuanced approximations of both self and others within the World-Model (Dunbar, 1998). *This social complexity is a major driver for the evolution of larger brains and **theory of mind** — the ability to attribute mental states to others (Tomasello, 1999).*

The development of consciousness, with its intricate interplay of ISM, World-Model, and Qualia, is not an accidental byproduct of biological complexity. It is, in the UAF framework, the most efficient and powerful mechanism that emerges from learning what reality approximately is. A conscious system, capable of building and refining its own approximations, can adapt to novel situations, learn from experience, and make flexible, goal-directed decisions far more effectively than a purely reflexive or hard-wired one (Godfrey-Smith, 2016), but it also inevitably approaches the truth that there is something it is like to be that system.

The primitive shared brain structures found in most mammals — such as the limbic system (involved in emotion and memory), the brainstem (regulating basic survival functions), and early cortical areas (for sensory processing) — are the biological foundations upon which these necessary approximations are built. *These structures represent a “triune brain” (MacLean, 1990) in a simplified sense, with older, more reflexive systems providing the bedrock for newer, more flexible cognitive capacities.* These structures, honed by millions of years of evolutionary pressure, represent the brain’s earliest attempts to ensure survival and reproduction. The very architecture of the mammalian brain is a testament to the force of Skin in the Game, pushing for the emergence of consciousness as the ultimate survival tool.

Chapter 30: The Architecture of Biological Qualia: Insights from Cognitive Science.

If **Qualia** are the brain’s “simplified truths” and “phenomenal flavors,” as **Useful Approximations Framework (UAF)** posits, then how are these subjective experiences instantiated and processed within the biological machinery of the brain? This chapter deepens the discussion on qualia by drawing insights from cognitive science and neuroscience, exploring how sensory modalities, internal representations, and neurological processing give rise to the specific, undeniable “feel” of biological consciousness.

Recall from **Chapter 8** that qualia serve two critical functional purposes: providing **Subjective Closure** and driving **Causal Efficacy**. They are the optimal compression of complex information into a directly usable, self-validating signal.

Consider the process of **color perception**. When light hits the retina, specialized photoreceptor cells (rods and cones) respond to different wavelengths. This is the initial, raw sensory input, part of the **Underlying Computational System (UCS)**. However, the “redness” we experience is not merely the wavelength of light. Instead, the signals from these photoreceptors are processed through multiple layers of neural networks in the visual cortex. These networks extract patterns, compare signals, and ultimately construct a simplified, internal representation (Zeki, 1993; Livingstone & Hubel, 1988). The quale of “red” is the brain’s unique, low-dimensional approximation of this underlying complexity. It’s a specific “phenomenal flavor” that is sufficiently accurate to successfully predict human behavior in everyday encounters — like identifying a ripe apple or recognizing a stop sign — without needing to process the infinite details of photon interactions.

Similarly, the experience of **pain** is a prime example of biological qualia. When tissue damage occurs, nociceptors (pain receptors) send electrical signals up the spinal cord to the brain. These signals activate a complex network of brain regions, including the thalamus, somatosensory cortex, insula, and anterior cingulate cortex (Craig, 2002). The “feeling” of pain is the brain’s integrated, simplified approximation of this vast neural activity and the underlying tissue damage. It’s a powerful, urgent quale that provides immediate **Subjective Closure** (you don’t need to interpret *why* it hurts; it just *does*) and drives **Causal Efficacy** (you immediately withdraw your hand). This low-dimensional approximation of “pain” is far more efficient for survival than processing the raw data of cellular damage. *This aligns with the Gate Control Theory of Pain, which posits that pain signals are modulated and filtered by the nervous system before reaching conscious awareness, emphasizing the brain’s active construction of the pain experience (Melzack & Wall, 1965).*

Insights from cognitive science, particularly in the field of **predictive coding**, further illuminate the architecture of biological qualia. The brain is constantly generating predictions about what it expects to perceive, and qualia arise when these predictions are either confirmed or significantly violated (Friston, 2010; Seth, 2021). The “surprise” or “prediction error” (Chapter 12) that results from unexpected sensory input can generate particularly vivid qualia, compelling the brain to update its **World-Model** and **Internal Self-Model (ISM)**. For example, the sudden, jarring quale of an unexpected loud noise forces an immediate update to your World-Model, signaling potential danger. *This suggests that qualia are not just passive readouts but active signals of informational salience, highlighting what is most important for the system to attend to and learn from (Hohwy, 2013).*

The brain’s architecture also demonstrates how qualia are deeply intertwined with our **Internal Self-Model**. Interoception, the sense of the physiological condition of the body, provides continuous input about our internal states—hunger, thirst, fatigue, heart rate. These internal signals are processed and often manifest as qualia (e.g., the dull ache of hunger, the sharp pang of thirst). These qualia are crucial for updating the ISM, allowing the brain to maintain a coherent, approximate understanding of its own

body and its needs, which in turn drives behaviors to maintain **Skin in the Game** (Damasio, 1999; Craig, 2002).

Furthermore, the brain’s capacity for **emotional qualia** (joy, sadness, anger, fear) highlights their role as powerful, compressed signals. These emotions are not just abstract concepts; they are felt experiences that provide immediate, simplified truths about our internal state in relation to our environment (Barrett, 2017). They guide our decisions, influence our social interactions, and motivate our actions, all as part of the brain’s persistent drive for **Imperative for Coherence & Agency**. *The somatic marker hypothesis* (Damasio, 1994) further suggests that these emotional qualia, or “somatic markers,” are crucial for rational decision-making, providing rapid, pre-conscious evaluations of potential outcomes.

In essence, the architecture of biological qualia is a testament to the brain’s genius in creating functionally indispensable internal simplified approximations of true reality. They are the low-dimensional, high-impact summaries of complex underlying processes, sufficiently accurate to successfully predict and guide human behavior in everyday encounters. Qualia are not an epiphenomenon; they are the very fabric of our subjective experience, computationally necessary for perception, action, and the continuous refinement of our conscious approximation of reality.

Chapter 31: Mental Illness as a Failure of Functional Fiction: A UAF Perspective.

If consciousness, as defined by **Useful Approximations Framework (UAF)**, is a “necessary functional fiction” — a system’s asymptotic best simplified approximation of what it is like to be an information processing system interacting with the universe — then what happens when this intricate system of approximation breaks down? This crucial chapter extends UAF’s explanatory power to mental illnesses, proposing that various mental health conditions can be understood as “**maladaptive functional fictions**” or, more simply, “**failed approximations of reality**.”

In a healthy mind, the **Internal Self-Model (ISM)**, **World-Model**, and **Qualia** work in concert, constantly refined through **Prediction Error Minimization (PEM)**, to provide a coherent, useful, and adaptive approximation of reality. This allows the individual to navigate their environment, maintain **Skin in the Game**, and achieve **Imperative for Coherence & Agency**. However, in mental illness, this delicate balance is disrupted. The brain forms massively broken representations or models of reality, which cause the person to behave irrationally, often to their own detriment.

Consider **psychosis**, particularly conditions like schizophrenia. Here, the brain’s World-Model and ISM generate approximations that deviate significantly from shared reality. Delusions are instances where the World-Model forms a “truth” that is not supported by external evidence, yet the system holds onto it with absolute certainty. Hallucinations are cases where the brain generates sensory qualia (e.g., voices, visions) with abnormally large error compared to the external input. Without access to the raw signals, the consciousness treats these internally generated “simplified truths” as the external reality (Frith, 1992). The individual is unable to learn or correct these models due to some underlying belief that is too scary or wonderful to change. As the underlying belief is held intact, the person needs to form complex explanations to support and protect it, leading to a cascade of further maladaptive approximations. *This process can be understood as a failure of metacognition — the ability to reflect on and evaluate one’s own thoughts and perceptions — leading to an inability to distinguish internal models from external reality* (Corlett et al., 2010). The system’s PEM mechanism, instead of correcting errors towards a shared reality, becomes trapped in a loop that reinforces the internal, distorted fiction.

Anxiety disorders can be understood as a failure in the predictive aspect of the World-Model and ISM, often coupled with dysfunctional qualia. The system constantly predicts threat or danger, even in safe environments (Barlow, 2002). The “fear” qualia (Chapter 8), which should be a high-bandwidth signal for actual danger, becomes overactive or miscalibrated, providing false “simplified truths” of threat (LeDoux, 1996). This leads to persistent prediction errors about safety, and the system’s attempts to minimize these errors result in maladaptive behaviors like avoidance or hyper-vigilance, further reinforcing the distorted World-Model. The individual’s **Subjective Closure** becomes trapped in a loop of perceived threat, even when objective reality offers no such evidence. *This persistent threat prediction often involves an overactive amygdala and impaired prefrontal cortex regulation, leading to a biased processing of ambiguous stimuli* (Etkin & Wager, 2007).

Depression, from a UAF perspective, can be seen as a profound failure in the system’s ability to

generate functionally useful approximations related to reward, motivation, and self-worth. The World-Model might become overly pessimistic, predicting negative outcomes regardless of effort. The ISM might form a severely diminished or negative approximation of the self, leading to feelings of worthlessness or hopelessness (Beck, 1967). Qualia related to pleasure or motivation become muted or absent, failing to provide the necessary “simplified truths” that drive engagement and goal pursuit (Rolls, 2000). The system loses its **Skin in the Game** for positive outcomes, leading to a state of **learned helplessness** (Seligman, 1975) where PEM struggles to find pathways to reduce prediction errors related to well-being. The individual’s capacity for **Causal Efficacy** is severely impaired, as the internal models no longer provide compelling reasons to act.

This perspective offers compelling real-world evidence for the consequences when consciousness’s essential mechanisms falter. Mental illnesses are not merely “chemical imbalances” (though neurochemistry plays a role in the UCS); they are profound dysfunctions in the brain’s ability to construct and maintain a coherent, adaptive, and useful “functional fiction” of reality and self. The brain forms massively broken representations or models of reality which cause the person to behave irrationally. The person is unable to learn or correct these models due to some underlying belief that is too scary or wonderful to change. As the underlying belief is held intact, the person needs to form complex explanations to support and protect it, leading to a cascade of further maladaptive approximations. *This resistance to updating, even in the face of contradictory evidence, highlights the **epistemic rigidity** that can characterize mental illness, where the “functional fiction” becomes entrenched and self-reinforcing (Maher, 1988).*

Understanding mental illness through the lens of UAF provides a powerful framework for both diagnosis and treatment. Therapies like **Cognitive Behavioral Therapy (CBT)** (Beck, 1967), for instance, can be seen as attempts to help individuals identify and correct these “failed approximations”—to challenge distorted World-Models and ISMs, and to re-calibrate the generation of maladaptive qualia, thereby guiding the system back towards a more useful and adaptive functional fiction. *This involves explicitly targeting the **prediction errors** that sustain maladaptive beliefs and behaviors, helping the brain to build more accurate and flexible models of self and world (Clark, 2013).* It underscores that the “truth” we experience is always an approximation, and when that approximation breaks down, the consequences are profoundly real.

*Read in the idiom of Chapter 9.5, mental illness is a node running persistently with $\mathcal{R} < 1$: $\mathcal{D}_{KL}$ is high, the system cannot lower it, Γ accumulates, and Φ (the cellular and cognitive substrate) is consumed in compensation. **Part IV-B**, which follows, treats the same accounting in agents whose $\mathcal{R}$ is at or above unity — and shows that the behavioural distinction we call “mental health” reduces, at the level of physics, to the same inequality that has organised every previous chapter of this book.*

Key References Cited (*Harvard Style, Alphabetical*)

- Barlow, D.H. (2002) *Anxiety and Its Disorders: The Nature and Treatment of Anxiety and Panic*, 2nd ed. Guilford Press.
- Barsalou, L.W. (2008) ‘Grounded Cognition’, *Annual Review of Psychology*, 59, pp. 617–645.
- Barrett, L.F. (2017) *How Emotions Are Made: The Secret Life of the Brain*. Houghton Mifflin Harcourt.
- Beck, A.T. (1967) *Depression: Causes and Treatment*. University of Pennsylvania Press.
- Buss, D.M. (1994) *The Evolution of Desire: Strategies of Human Mating*. Basic Books.
- Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chaitin, G. (2005) *Meta Maths: The Quest for Omega*. Vintage.
- Clark, A. (2013) ‘Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science’, *Behavioral and Brain Sciences*, 36(3), pp. 181–204.
- Clark, A. (2016) *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.
- Corlett, P.R. et al. (2010) ‘Prediction Errors and Psychosis: Recent Advances and Implications for Early Intervention’, *Schizophrenia Bulletin*, 36(6), pp. 1041–1048.
- Craig, A.D. (2002) ‘How Do You Feel? Interoception: The Sense of the Physiological Condition of the Body’, *Nature Reviews Neuroscience*, 3(8), pp. 655–666.
- Damasio, A. (1994) *Descartes’ Error: Emotion, Reason, and the Human Brain*. G.P. Putnam’s

Sons.

- **Damasio, A.** (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt Brace.
- **Darwin, C.** (1859) *On the Origin of Species*. John Murray.
- **Davidson, E.H.** (2006) *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution*. Academic Press.
- **Dawkins, R.** (1976) *The Selfish Gene*. Oxford University Press.
- **Dennett, D.** (1991) *Consciousness Explained*. Little, Brown and Company.
- **Dunbar, R.I.M.** (1998) 'The Social Brain Hypothesis', *Evolutionary Anthropology: Issues, News, and Reviews*, 6(5), pp. 178–190.
- **Etkin, A. and Wager, T.D.** (2007) 'Functional Neuroimaging of Anxiety: A Meta-Analysis of Emotional Processing in PTSD, Social Anxiety Disorder, and Specific Phobia', *American Journal of Psychiatry*, 164(10), pp. 1476–1488.
- **Fahle, M.** (2005) 'Perceptual Learning: A Review of Mechanisms and Clinical Applications', *Progress in Neurobiology*, 77(1–2), pp. 14–34.
- **Felleman, D.J. and Van Essen, D.C.** (1991) 'Distributed Hierarchical Processing in the Primate Cerebral Cortex', *Cerebral Cortex*, 1(1), pp. 1–47.
- **Fields, R.D.** (2009) 'The Other Half of the Brain', *Scientific American*, 300(4), pp. 54–61.
- **Friston, K.** (2010) 'The Free-Energy Principle: A Unified Brain Theory?', *Nature Reviews Neuroscience*, 11(2), pp. 127–138.
- **Frith, C.D.** (1992) *The Cognitive Neuropsychology of Schizophrenia*. Lawrence Erlbaum Associates.
- **Godfrey-Smith, P.** (2016) *Other Minds: The Octopus and the Evolution of Intelligent Life*. HarperCollins.
- **Herculano-Houzel, S.** (2009) 'The Human Brain in Numbers: A Linearly Scaled-Up Primate Brain', *Frontiers in Human Neuroscience*, 3(31).
- **Hofstadter, D.** (1979) *Gödel, Escher, Bach: An Eternal Golden Braid*. Basic Books.
- **Hohwy, J.** (2013) *The Predictive Mind*. Oxford University Press.
- **Kahneman, D.** (2011) *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- **LeDoux, J.** (1996) *The Emotional Brain*. Simon & Schuster.
- **Livingstone, M. and Hubel, D.** (1988) 'Segregation of Form, Color, Movement, and Depth: Anatomy, Physiology, and Perception', *Science*, 240(4853), pp. 740–749.
- **MacLean, P.D.** (1990) *The Triune Brain in Evolution: Role in Paleocerebral Functions*. Plenum Press.
- **Maher, B.A.** (1988) 'Anomalous Experience and the Psychopathology of Delusions', *Annals of the New York Academy of Sciences*, 563(1), pp. 103–119.
- **Marr, D.** (1982) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press.
- **Melzack, R. and Wall, P.D.** (1965) 'Pain Mechanisms: A New Theory', *Science*, 150(3699), pp. 971–979.
- **Metzinger, T.** (2003) *Being No One: The Self-Model Theory of Subjectivity*. MIT Press.
- **Metzinger, T.** (2009) *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. Basic Books.
- **Rolls, E.T.** (2000) 'The Brain and Emotion: The Neurobiology of Affective States', *Oxford University Press*.
- **Seligman, M.E.P.** (1975) *Helplessness: On Depression, Development, and Death*. W.H. Freeman.
- **Seth, A.** (2021) *Being You: A New Science of Consciousness*. Dutton.
- **Shettleworth, S.J.** (2010) *Cognition, Evolution, and Behavior*, 2nd ed. Oxford University Press.
- **Sterelny, K.** (2003) *Thought in a Hostile World: The Evolution of Human Cognition*. Blackwell.
- **Taleb, N.N.** (2018) *Skin in the Game: Hidden Asymmetries in Daily Life*. Random House.
- **Tomasello, M.** (1999) *The Cultural Origins of Human Cognition*. Harvard University Press.
- **Treisman, A.** (1996) 'The Binding Problem', *Current Opinion in Neurobiology*, 6(2), pp. 171–178.
- **Zeki, S.** (1993) *A Vision of the Brain*. Blackwell Scientific Publications.