

Temporal Predictions and Distributed Minds:

A Comparative HTM-Based Framework for Consciousness in Humans and Octopuses

Alex Horovitz

ahorovitz@mail.sfsu.edu

April 30, 2025

San Francisco State University
Department of Philosophy

Abstract

[Insert abstract here... or add `\input{abstract.tex}` to include from a file.]

1 Introduction

The mystery of consciousness—what it is, how it arises, and in whom or what it resides—remains one of the most profound and enduring questions in neuroscience and philosophy. Despite significant advances in neuroimaging and computational modeling, the so-called “hard problem” of consciousness, famously articulated by David Chalmers [Chalmers, 1996], persists: how and why do physical processes in the brain give rise to subjective experience? In the search for answers, a major focus of research has shifted toward identifying the *neural correlates of consciousness* (NCCs), or the specific patterns of brain activity that systematically accompany conscious awareness. Yet, while correlates are helpful, they do not explain the *mechanism*—*how* consciousness emerges from neural dynamics.

A growing body of work in theoretical neuroscience suggests that consciousness may be deeply tied to *predictive processing*. According to this framework, the brain is fundamentally a prediction machine: it constructs internal models of the world and continuously updates them by minimizing the error between its predictions and incoming sensory input. Theories like Karl Friston’s free energy principle [Friston, 2010] and Andy Clark’s hierarchical predictive coding [Clark, 2016] present the mind as a system of nested predictors, each contributing to perception, action, and cognition through the minimization of surprise. Within this conceptual landscape, consciousness is increasingly seen not as a passive reflection of sensory input, but as an *active inference* process—where perception is the brain’s best guess about what is out there, given past experience and current context.

A particularly elegant instantiation of the predictive processing paradigm is found in the *Hierarchical Temporal Memory* (HTM) theory developed by Jeff Hawkins [Hawkins and Blakeslee, 2004], drawing directly from Vernon Mountcastle’s seminal insight that the neocortex operates through a repeated, canonical computational unit—the cortical column [Mountcastle, 1978]. In HTM, each cortical column learns sequences of patterns over time, encodes them sparsely, and makes predictions about what will happen next. These predictions are constantly tested against sensory input, and mismatches (prediction errors) are used to refine the system. The neocortex, in this view, is a *temporally organized hierarchy of predictors*—and it is this structure that gives rise to a unified, dynamic, and context-sensitive conscious experience.

Consciousness, then, may not be a “thing” localized in a specific region of the brain, but an emergent property of hierarchical temporal memory systems resolving uncertainty across time.

While much of the neuroscientific and philosophical literature on consciousness remains centered on human brain architecture—particularly the neocortex—there is increasing interest in non-mammalian species that exhibit complex behavior despite having radically different neural organizations. The octopus, in particular, has become a compelling case study. As a cephalopod mollusk, the octopus possesses a large and intricate nervous system, yet one that is *fundamentally decentralized*: more than half of its neurons reside in its arms, which can process information and even execute goal-directed behavior independently of the central brain [Carls-Diamante, 2022]. Octopus cognition appears to operate in a *distributed* fashion, with limbs acting semi-autonomously while still contributing to the behavior of the whole organism. Behavioral research suggests rich perceptual experience, curiosity, problem-solving, and affect-based decision making [Mather, 2021]. This unique arrangement raises an intriguing question: **Can a system without a centralized cortical-like architecture still give rise to conscious experience?**

In this paper, I explore whether consciousness can emerge from predictive architectures that are not centrally organized. Building on HTM theory, I argue that the octopus nervous system may represent a *distributed implementation of temporal memory*, where each neural module (such as an arm ganglion) operates as a localized predictor, and coherence emerges from the interaction among these modules. By comparing the human and octopus systems through the lens of temporally-organized predictive processing, I aim to assess whether centralized integration (e.g., via the corpus callosum or global neuronal workspace) is necessary for consciousness, or whether embodied, distributed predictors can achieve a similar unification of experience.

Ultimately, this comparative, cross-species approach to the neural basis of consciousness challenges anthropocentric assumptions and invites a broader framework—one in which consciousness is not limited to brains that resemble our own, but may instead be understood as a function of temporally structured inference systems, whether centralized in the neocortex or distributed across limbs in an octopus.

2 Theoretical Background:

Predictive Processing and Hierarchical Temporal Memory

One of the most influential contemporary frameworks for understanding cognition and consciousness is the predictive processing model. This view holds that the brain is fundamentally a prediction machine, constantly generating hypotheses about sensory input and updating its internal models to minimize the difference between expectation and reality [Friston, 2010, Clark, 2016]. According to predictive processing theories, perception, action, and thought are unified under a single overarching principle: the minimization of prediction error, or “surprise,” in an organism’s interactions with the world.

Under this model, the brain maintains a generative model of the environment, issuing top-down predictions about incoming sensory data at every level of its processing hierarchy. Lower-level sensory cortices respond primarily to deviations from these predictions—known as prediction errors—which are fed back upward to adjust and refine the internal model [Rao and Ballard, 1999]. Thus, rather than passively receiving information, the brain is actively inferring the most probable causes of sensory inputs. Higher-order processes are concerned not only with interpreting signals, but with reducing uncertainty over time, constantly fine-tuning expectations to create coherent perceptual experiences.

Central to this framework is the notion of a hierarchical organization of the cortex. Each layer of the hierarchy is responsible for predicting the activity of the layer below it, with ascending signals carrying only error terms [Friston, 2005]. The predictive processing model aligns naturally with notions of sparse coding and hierarchical temporal structure observed in cortical organization, offering a plausible mechanism by which conscious awareness emerges through layered, dynamic inference.

A particularly compelling instantiation of this idea at the algorithmic level is found in Hierarchical Temporal Memory (HTM), a theory proposed by Jeff Hawkins [Hawkins and Blakeslee, 2004]. HTM takes inspiration from the anatomical structure of the neocortex, particularly Vernon Mountcastle’s insight that the cortical column is a universal, repeating unit of computation throughout the cortex [Mountcastle, 1978]. In HTM, each cortical column learns sequences of patterns over time, stores them using sparse distributed representations, and predicts

future inputs based on temporal context.

Unlike many traditional machine learning models that emphasize static feature recognition, HTM prioritizes the learning of transitions and temporal dependencies. Each cell in an HTM model not only activates in response to a particular input but does so in a context-dependent manner, remembering the sequence of prior activations that led to the current state. This architecture enables the system to make robust predictions about upcoming events even under conditions of uncertainty or noise.

The emphasis on sparsity is crucial. Sparse distributed representations (SDRs) allow HTM systems to encode information with high capacity and noise tolerance while ensuring that overlapping representations maintain semantic similarity [Ahmad and Hawkins, 2015]. This property mirrors empirical findings from neuroscience, which show that real cortical representations are highly sparse, with only a small percentage of neurons active at any given time [Olshausen and Field, 1996].

Further, HTM posits a deep interaction between spatial and temporal learning. While sensory cortices encode spatial patterns, the temporal memory of HTM allows the system to encode sequences of these spatial patterns, effectively linking space and time into a single predictive framework. This approach mirrors the way in which the human brain is thought to integrate temporally evolving information to form coherent streams of consciousness [Seth, 2015].

In predictive processing frameworks broadly—and HTM specifically—consciousness is not seen as a passive byproduct of sensory input but as an active, ongoing process of model refinement. As prediction errors accumulate and are resolved, the system continuously updates its internal representations, effectively constructing a subjective reality in real time. This view contrasts sharply with classical computational theories of mind, which often treat perception as a direct pipeline from sensation to cognition.

In sum, predictive processing and HTM converge on several key points: the brain as a hierarchical, predictive engine; the fundamental role of temporal sequence learning; the importance of sparse, distributed codes; and the view of consciousness as emerging from dynamic, hierarchical model optimization. This theoretical background provides a powerful lens for interpreting consciousness not as an intrinsic property of biological tissue per se, but

as a structural and functional consequence of organized prediction across time and space.

3 The HTM Model of Consciousness

Hierarchical Temporal Memory (HTM) offers a biologically inspired framework for understanding intelligence and consciousness based on the structural and functional organization of the neocortex. Developed by Jeff Hawkins and colleagues, HTM builds directly upon Vernon Mountcastle’s hypothesis that the cortex is composed of uniform, repeating computational units known as cortical columns [Hawkins and Blakeslee, 2004, Mountcastle, 1978]. While Mountcastle initially described this architecture anatomically, HTM provides a computational model that specifies how these columns might work together to support predictive, temporally extended cognition.

At the core of HTM is the concept of *temporal memory*. Each cortical column in the HTM framework learns sequences of spatial patterns over time, allowing it to predict future inputs based on contextual history. Importantly, HTM systems employ *sparse distributed representations* (SDRs), where only a small subset of neurons is active at any given time, enabling efficient, robust encoding of information with high noise tolerance [Ahmad and Hawkins, 2015, Hawkins and Ahmad, 2016].

In the HTM model, neurons are modeled not just as simple threshold units, but as multi-dimensional predictors. Each neuron possesses multiple dendritic segments that recognize different patterns of input; activation of a predictive segment can depolarize the neuron, putting it into a predictive state even before the input arrives. This allows the system to anticipate sequences and to distinguish between different temporal contexts [Hawkins and Ahmad, 2016].

Prediction is therefore not merely about static classification, but about learning the transitions between states. As a result, HTM networks excel in environments where understanding change over time is critical, offering a natural foundation for cognitive processes like perception, motor planning, and higher-order thought. Temporal memory creates a rolling “expectation” that is constantly tested against sensory input, facilitating a dynamic interaction between the organism and its environment [Hawkins, 2017].

Consciousness as Temporal Predictive Integration

In the context of consciousness, HTM provides a compelling model by suggesting that conscious experience emerges from the integration of hierarchical, temporally organized predictions. Each layer in the cortical hierarchy learns progressively more abstract sequences: lower layers predict simple sensory transitions, while higher layers predict more abstract and temporally extended events.

This hierarchical prediction structure mirrors how subjective experience is organized. Phenomenologically, conscious experience is continuous and predictive: we anticipate the movement of objects, the flow of speech, and even our own intentions [Clark, 2016, Seth, 2015]. HTM models explain this by positing that higher layers in the hierarchy compress and integrate sequences over longer timescales, giving rise to coherent, temporally structured awareness.

Furthermore, the concept of *attention* can be naturally incorporated into the HTM framework. Attention in HTM-like systems can be viewed as the modulation of prediction precision—allocating more resources to uncertain or surprising inputs, consistent with predictive coding models of attention [Feldman and Friston, 2010]. Regions receiving surprising input can dynamically recruit more processing resources, shifting the focus of conscious awareness.

Selfhood and the Predictive Model of the Self

Another dimension of consciousness is selfhood—the experience of being an agent embedded in an environment. Recent research suggests that the brain maintains an internal predictive model of the body and its interactions with the world [Apps and Tsakiris, 2015]. In HTM, such a model could arise naturally as part of hierarchical temporal learning: cortical regions responsible for sensorimotor integration would predict the consequences of self-generated actions, forming an evolving, embodied model of the self.

The neocortex’s ability to generate counterfactual predictions—anticipating the outcomes of imagined actions—is crucial for goal-directed behavior and planning. In the HTM framework, such counterfactual prediction is a natural extension of temporal sequence memory, where imagined futures are treated as internally simulated sequences that can be compared

against predicted outcomes.

Summary: Consciousness as Predictive Compression

HTM thus suggests that consciousness is not a passive mirroring of the world but an active, compressed summary of hierarchical, temporally organized predictions. Conscious experience arises as the brain continuously updates its internal models, integrating sensory, motor, and internal states into a coherent predictive whole. This model aligns with and extends predictive processing theories by emphasizing the importance of sequence learning, sparse representation, and hierarchical abstraction.

In contrast to static snapshot views of perception, HTM models consciousness as an intrinsically dynamic process: a fluid, temporally structured dance of predictions and corrections. This dynamic view not only accounts for core phenomenological aspects of consciousness but also offers a computationally plausible mechanism rooted in known neuroanatomical and physiological principles.

4 Octopus Neuroanatomy and Cognition

The octopus presents a profound challenge to traditional human-centric models of cognition and consciousness. As a member of the cephalopod mollusks, octopuses exhibit a unique nervous system architecture that diverges radically from that of vertebrates. Unlike the centralized organization of the mammalian brain, the octopus nervous system is *highly decentralized*, with the majority of its neurons located outside of the brain, distributed throughout its eight flexible arms [Hochner, 2012, Godfrey-Smith, 2016].

The octopus brain is composed of approximately 500 million neurons, a number comparable to that of some small mammals. However, an estimated two-thirds of these neurons reside in the nerve cords of the arms, organized into ganglia that operate with significant autonomy [Godfrey-Smith, 2016]. Each arm is capable of executing complex motor actions such as reaching, grasping, and exploratory behavior with minimal input from the central brain [Zylinski et al., 2013]. Studies have shown that arms can perform reflexive withdrawal, navigate mazes, and adapt their movement strategies even when severed from the central

brain [Gutnick et al., 2011].

This distributed control structure suggests that the octopus operates as a network of semi-independent agents coordinated by a supervisory central brain, rather than as a hierarchical system of command and control. The arm ganglia appear to integrate local sensory information and execute motor commands directly, only occasionally relaying high-level summaries to the brain for further integration. From an architectural perspective, this decentralization resembles a form of *embodied computation* in which intelligence is not confined to the head but extends throughout the body.

Octopus Behavior and Evidence of Complex Cognition

Despite their unusual neural architecture, octopuses demonstrate behaviors typically associated with higher-order cognitive processing. They exhibit sophisticated problem-solving abilities, such as opening jars, escaping enclosures, and navigating complex mazes [Mather, 2008]. Observations of play behavior—manipulating objects without clear survival value—suggest an internal motivational structure extending beyond simple stimulus-response mechanisms [Kuba et al., 2006].

Moreover, octopuses display remarkable flexibility in their behavioral strategies, often using trial-and-error learning to adapt to novel situations. They show evidence of episodic-like memory, anticipating future events based on past experiences, and modifying their actions accordingly [Fidalgo-Mathies and Sumbre, 2021]. These capacities imply the existence of an internal model of the environment that supports planning and expectation, traits commonly associated with conscious processing.

Research into octopus perception further supports the notion of complex subjective experience. Octopuses possess sophisticated visual systems capable of polarization sensitivity and dynamic camouflaging based on visual context [Kelman et al., 2006]. Their ability to modify body texture and coloration implies an ongoing integration of sensory input with motor control in real time, requiring a detailed, moment-to-moment representation of both body and environment.

Distributed Cognition and Consciousness in the Octopus

The decentralized structure of the octopus nervous system raises important questions about the organization of consciousness. Traditional models of consciousness often presume the necessity of a centralized integrating hub, such as the mammalian thalamocortical system. However, in the octopus, integration appears to occur through the interaction of multiple semi-autonomous modules, each capable of learning, predicting, and adapting independently.

Sidney Carls-Diamante has argued that octopus cognition challenges classical conceptions of unitary conscious experience, suggesting that conscious states may be distributed and modular rather than unified [Carls-Diamante, 2022]. If each arm processes information semi-independently while still contributing to overall behavior, it is plausible that conscious experience in the octopus is similarly partitioned, either existing in a fragmented form or integrated through non-traditional mechanisms.

Peter Godfrey-Smith has proposed that consciousness in the octopus may arise not from centralized control but from dynamic coordination among multiple neural systems [Godfrey-Smith, 2016]. In this view, consciousness could emerge from the negotiation between local predictors—each with partial access to the environment and body—producing a unified experience through continuous interaction rather than centralized representation.

Implications for Theories of Mind

The case of the octopus suggests that consciousness need not be confined to architectures resembling the vertebrate brain. It opens the possibility that conscious experience could arise from distributed, embodied predictive systems rather than from a monolithic, centralized processor. This distributed model aligns well with recent trends in embodied cognition and predictive processing, which emphasize the tight coupling of perception, action, and environmental context [Clark, 2016].

The octopus challenges researchers to rethink fundamental assumptions about the neural correlates of consciousness, suggesting that temporally organized prediction, rather than anatomical centralization, may be the critical feature underlying subjective experience. This shift holds profound implications for comparative neuroscience, philosophy of mind, and

artificial intelligence, broadening the landscape of what kinds of systems might be capable of consciousness.

References

- [Ahmad and Hawkins, 2015] Ahmad, S. and Hawkins, J. (2015). Properties of sparse distributed representations and their application to hierarchical temporal memory. *arXiv preprint*, arXiv:1503.07469.
- [Apps and Tsakiris, 2015] Apps, M. A. J. and Tsakiris, M. (2015). Prediction and the brain: How predictive coding explains attention, social cognition, and the self. *Brain Research*, 1626:158–176.
- [Carls-Diamante, 2022] Carls-Diamante, S. (2022). Where is it like to be an octopus? *Frontiers in Systems Neuroscience*, 16:840022.
- [Chalmers, 1996] Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, New York.
- [Clark, 2016] Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press, Oxford.
- [Feldman and Friston, 2010] Feldman, H. and Friston, K. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4:215.
- [Fidalgo-Mathies and Sumbre, 2021] Fidalgo-Mathies, J. and Sumbre, G. (2021). Octopus episodic-like memory: Behavioral evidence and the implications for consciousness. *Frontiers in Systems Neuroscience*, 15:682487.
- [Friston, 2005] Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456):815–836.
- [Friston, 2010] Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138.

- [Godfrey-Smith, 2016] Godfrey-Smith, P. (2016). *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. Farrar, Straus and Giroux, New York.
- [Gutnick et al., 2011] Gutnick, T., Byrne, R. A., Hochner, B., and Kuba, M. J. (2011). Tactile discrimination in the octopus: Fine control of octopus arm movements. *Journal of Comparative Physiology A*, 197:903–912.
- [Hawkins, 2017] Hawkins, J. (2017). A thousand brains: A new theory of intelligence. *arXiv preprint*, arXiv:1707.06208.
- [Hawkins and Ahmad, 2016] Hawkins, J. and Ahmad, S. (2016). Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Frontiers in Neural Circuits*, 10:23.
- [Hawkins and Blakeslee, 2004] Hawkins, J. and Blakeslee, S. (2004). *On Intelligence*. St. Martin’s Press, New York.
- [Hochner, 2012] Hochner, B. (2012). An embodied view of octopus neurobiology. *Current Biology*, 22(20):R887–R892.
- [Kelman et al., 2006] Kelman, E. J., Osorio, D., and Baddeley, R. J. (2006). Polarization signals and the response of marine organisms. *Marine and Freshwater Behaviour and Physiology*, 39(1):1–14.
- [Kuba et al., 2006] Kuba, M. J., Byrne, R. A., Meisel, D. V., and Mather, J. A. (2006). Do octopuses play? *Journal of Comparative Psychology*, 120(3):184–190.
- [Mather, 2008] Mather, J. A. (2008). Cephalopod consciousness: behavioural evidence. *Consciousness and Cognition*, 17(1):37–48.
- [Mather, 2021] Mather, J. A. (2021). Octopus consciousness: The role of perceptual richness. *NeuroSci*, 2(3):276–290.
- [Mountcastle, 1978] Mountcastle, V. B. (1978). An organizing principle for cerebral function: The unit model and the distributed system. In Edelman, G. M. and Mountcastle, V. B., editors, *The Mindful Brain*, pages 7–50. MIT Press, Cambridge, MA.

- [Olshausen and Field, 1996] Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609.
- [Rao and Ballard, 1999] Rao, R. P. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87.
- [Seth, 2015] Seth, A. K. (2015). Prediction, inference, and consciousness. *Consciousness and Cognition*, 36:1–12.
- [Zylinski et al., 2013] Zylinski, S., Osorio, D., and Shohet, A. J. (2013). Visual control of arm movement in the octopus. *Frontiers in Physiology*, 4:97.