

The Mind That Lies: Constructed Perception, the Inescapable Self-Model, and the Plurality of Control in Conscious Experience

Philosophy of Mind • Cognitive Neuroscience • Consciousness Studies

Mamidala Harsha Vardhan
ORCID: 0009-0000-7656-4353

*Mamidala Harshavardhan, Electronics and Communication Engineer, Embedded Firmware Engineer,
Saaradaa Learknowations Private Limited, Nagole 500068, Hyderabad, Telangana State.*

Abstract:

We take it for granted that we perceive the world as it is, that we stand outside our experience as its impartial witness, and that a single self authors our thoughts and choices. This paper argues that all three convictions are constructions of the brain rather than facts about reality. Drawing together predictive-processing accounts of perception, the self-model theory of subjectivity, and competition-based theories of consciousness, it advances three linked claims: first, that what we perceive is a generative model - a “controlled hallucination” - and not the world itself; second, that this model is inescapable by design, because its own constructed character is rendered transparent to us, so that knowing it is a model does not let us step outside it; and third, that the unified self who seems to perceive and decide is a narrative imposed after the fact upon many competing subsystems contending for control. The argument is anchored empirically in recent work showing that even the most basic structural feature of ordinary experience - the division into an inner self and an outer world - tracks a measurable difference in the brain’s intrinsic neural timescales, and that this difference collapses in advanced meditators who report non-dual awareness (Saketh et al., 2026). That the self–world boundary can be dissolved is the strongest available evidence that it was built. The mind does not lie maliciously; it lies usefully. But the gap between what we live and what is the case is real, permanent for most of us, and worth understanding.

Keywords:

consciousness; predictive processing; controlled hallucination; phenomenal self-model; naive realism; non-duality; intrinsic neural timescales; global workspace; the divided self.

1. Introduction:

The Question of the Real:

Open your eyes and the world is simply *there* - coloured, solid, and outside you - delivered to a self who looks out upon it from somewhere behind the face. Nothing in ordinary experience suggests that this scene is manufactured. It feels received, not made. Yet a century and a half of work in perception, philosophy of mind, and neuroscience converges on an uncomfortable conclusion: the felt directness of experience is precisely the part that is illusory. We do not see the world; we see a model of the world, built by the brain and presented to us with its seams hidden. The conviction that we have unmediated access to reality is not a perception of reality. It is a feature of the model.

This essay defends a strong reading of that conclusion - captured bluntly in the idea that **the mind lies to us**. It develops the claim in three stages, each grounded in an established research programme, and binds them to a recent empirical result that gives the whole argument an unusually concrete footing: the demonstration by Saketh, Northoff, and colleagues that the brain's division of experience into "inner" and "outer" is a tunable construction, indexed by neural timescales, that can be measurably dissolved (Saketh et al., 2026).

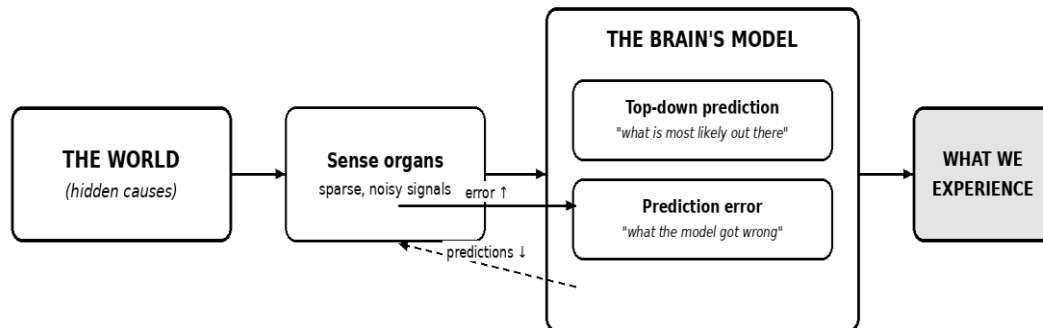
Thesis and Structure:

Three propositions organise what follows. **First**, what we perceive is a constructed model rather than the world itself; perception is controlled hallucination (Section 2). **Second**, we cannot escape the model, because the brain renders the model *transparent* - we look through it to a world that seems given, and merely knowing it is a model does not change how it appears (Section 4). **Third**, the unified self who seems to do the perceiving and the choosing is itself a construction: a narrative laid over many competing processes, none of which is the "I" (Section 5). Between the first and the later claims stands the keystone (Section 3): the boundary between self and world - the most basic frame of all experience - is not given but built, and the evidence is that it can be taken apart.

2. The Constructed World: Perception as Inference:

That perception is inference rather than transcription is among the oldest and best-supported ideas in the science of the mind. Helmholtz proposed in the nineteenth century that perception is a process of *unconscious inference*: the senses deliver signals that are sparse, noisy, and radically ambiguous, and the brain must guess at their most probable cause (von Helmholtz, 1867/1925). The image on the retina is two-dimensional, inverted, and riddled with gaps, yet we experience a stable, three-dimensional world. The surplus is supplied not by the eye but by the brain's expectations.

Modern theory has formalised this intuition. On the predictive-processing account, the brain is a hierarchical prediction machine: higher levels continuously generate predictions of the incoming signal, lower levels return only the *error* - the difference between prediction and input - and perception is the brain's best current hypothesis, the one that explains away the most error (Clark, 2013; Friston, 2010; Hohwy, 2013). The free-energy principle casts this as a single imperative to minimise long-run prediction error, unifying perception, learning, and action under one rule (Friston, 2010). On this view the flow of information is mostly top-down: what we see is dominated by what the brain expects to see, with the senses serving chiefly to correct and constrain the guess (Clark, 2016). Figure 1 sketches the loop.

Figure 1. Perception as construction: the predictive loop

We never perceive the world directly. The brain builds a best-guess model from prior expectations and corrects it with sensory error. What reaches awareness is the model - a "controlled hallucination" kept in check by, but never identical to, reality.

Fig. 1. Perception as a predictive loop. The brain's model - not the sensory stream - is what reaches awareness; sensory error constrains the model but never replaces it.

Anil Seth has given this picture its sharpest popular formulation, describing perception as a *controlled hallucination*: the same constructive machinery that produces dreams and hallucinations produces ordinary seeing, the only difference being that waking perception is reined in by sensory signals from the world (Seth, 2021). The continuity is the point. Perception and hallucination are not opposites; they are the same process under different degrees of sensory control.

"We're all hallucinating all the time, including right now. It's just that when we agree about our hallucinations, we call that reality."

- Anil Seth (2021)

The constructive account is not an interpretation imposed on otherwise neutral data; it is forced by the data. Visual illusions persist even when we know they are illusions - the two equal lines of the Müller-Lyer figure refuse to look equal - because the inference that produces them is encapsulated from belief (Gregory, 1997). The brain *fills in* the retinal blind spot with invented content we never notice. Colour, famously, is not a property of surfaces but a construction the brain assigns to wavelength; perceptual constancies hold an object's apparent size and colour steady as the raw signal swings wildly; and a single unchanging image, as in bistable figures, flips between rival interpretations while the input stays fixed. In each case experience tracks the brain's hypothesis, not the stimulus.

Evolution offers a reason the construction need not be accurate. Hoffman argues that natural selection favours perceptual systems tuned to *fitness*, not to truth: an organism that perceives reality veridically will, in general, be out-competed by one that perceives a simplified, action-guiding interface (Hoffman, 2019). On this reading our perceptions are less like a window onto reality than like the icons on a computer desktop - useful, manipulable symbols that conceal the machinery they stand for. The folder on the screen is neither the true shape nor the true location of the file; it is a fiction that lets us act. So, the argument runs, is the apple, the colour red, and the solid table.

3. The Self–World Boundary Is Built, Not Given:

If particular percepts are constructed, one might still suppose that the *frame* within which they appear - the basic division between an inner self and an outer world - is simply given, the bedrock on which all experience rests. This supposition is the deepest part of the illusion, and it is the part for which we now have direct neural evidence. The distinction between self and world is not the floor beneath the construction; it is itself constructed.

Contemplative traditions have asserted as much for millennia, describing *non-dual* states in which the boundary between subject and object dissolves. Josipovic frames duality as the structuring of experience into a subject–object dichotomy and, more broadly, into reified polarities - self and other, inner and outer, mind and body - while non-duality is experiencing and knowing without those divisions (Josipovic, 2019). For most of the history of the subject these reports could be dismissed as ineffable or merely metaphorical. Recent neuroscience has changed that.

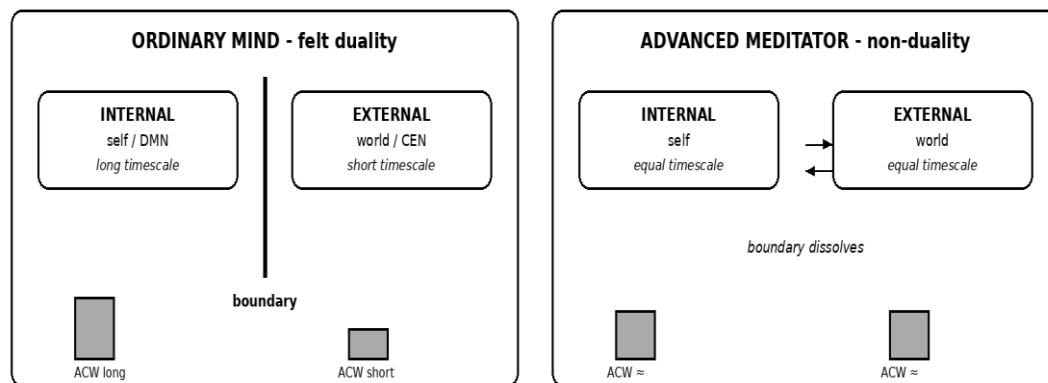
Saketh, Northoff, and colleagues operationalised non-duality and measured it (Saketh et al., 2026). Using EEG, they indexed the brain's *intrinsic neural timescales* - how long neural activity remains correlated with its own recent past, estimated by the autocorrelation window - during inward attention (breath-watching meditation) and outward attention (a visual oddball task). Intrinsic timescales are a basic property of cortical dynamics, longer in regions that integrate information over time and shorter in those that track fast external change (Golesorkhi et al., 2021). Intuitively, the measure asks how long the brain “holds on” to its own activity: a long timescale means the present moment is shaped heavily by what came just before - as when attention rests in an enduring sense of self - whereas a short timescale means activity turns over quickly to keep pace with a changing world. In ordinary participants these timescales were reliably longer during inward than outward attention - a measurable neural signature of the self–world split, which the authors term *neuro-temporal duality*. The inner-directed default-mode system and the outer-directed executive system (Raichle, 2015; Northoff & Bermpohl, 2004) operated on visibly different clocks.

The decisive finding concerns the advanced meditators. In them the difference largely vanished: their intrinsic timescales were *the same* whether attention was turned inward or outward - a state the authors call *neuro-temporal non-duality* - and across all participants a smaller inner–outer timescale difference predicted a stronger reported experience of non-duality, mediated specifically by equanimity (Saketh et al., 2026). The pattern was not an artefact of one small sample: the authors reproduced the absence of an inner–outer timescale difference in an independent group of expert meditators, and showed it could not be explained

away by differences in alpha-band power or in whether the eyes were open or closed. Figure 2 summarises the contrast.

Figure 2. The self-world boundary is built, not given

after Saketh, Northoff et al. (2026): *intrinsic neural timescales index the felt divide*



The most basic structure of ordinary experience - a self set against a world - tracks a difference in the brain's intrinsic timescales for inward vs. outward attention. In advanced meditators that difference collapses and the boundary is no longer felt. If even the self-world divide is a tunable construction, the partition is the brain's doing, not reality's.

Fig. 2. The self-world boundary as a neural construction. The felt divide between self and world tracks a difference in intrinsic neural timescales for inward versus outward attention; in advanced meditators the difference collapses and the boundary is no longer experienced (after Saketh et al., 2026).

The philosophical weight of this result is hard to overstate. The boundary between self and world is the most fundamental structural feature of ordinary consciousness - the precondition for there being an “I” to whom a “world” appears. To find that this boundary corresponds to a particular configuration of neural timescales, and that *dissolving the configuration dissolves the boundary*, is to find that the boundary is a construction of the brain and not a feature of reality. A partition that can be tuned, narrowed, and switched off by training is the brain's doing, not the world's. This is the empirical keystone of the present argument: the constructedness of experience is demonstrated, not merely inferred, at the very level one would have thought immune to it. The contemplative claim that the embodied mind has no fixed ground (Varela et al., 1991) here meets quantitative confirmation.

4. The Inescapable Model: Transparency and Naïve Realism:

If experience is constructed, why can we not simply notice the construction and look past it? The answer is that the brain hides its work. Metzinger calls the crucial property *phenomenal transparency*: we do not experience our model of the world *as* a model; we look straight through it to a world that seems immediately present, in just the way we look through a clean

pane of glass without seeing the glass (Metzinger, 2003). The representational vehicle is invisible; only its content is given. Transparency is why reality feels received rather than rendered, and it applies with special force to the model of the self: the self-model, too, is transparent, so that we identify with it completely and cannot experience ourselves as systems running a self-simulation. We live, in his phrase, inside an *ego tunnel* - a tunnel whose walls we mistake for the horizon (Metzinger, 2009).

The conscious self is the content of a model that cannot be recognised as a model. What we call "reality" is a virtual scene rendered so seamlessly that the rendering is never seen - only the world it depicts.

- after Metzinger (2003, 2009)

Crucially, the illusion is not dispelled by understanding it. This is what makes it a *design feature* rather than a mistake we might correct. The Müller-Lyer lines go on looking unequal after they are measured; the perceptual inference is cognitively impenetrable, sealed off from what we believe (Gregory, 1997). One can know, in full intellectual detail, that colour is a construction and that the self is a model, and the apple will still look simply red and the self will still feel simply real. Knowledge buys no exit. The brain's default is *naïve realism* - the unshakeable sense that we perceive things as they are - and the brain actively conceals the constructive labour that produces it (Frith, 2007). We are not given a model and a world to compare; we are given only the model, labelled "world."

Even the non-dual states of Section 3 are not an escape from this predicament so much as proof of its depth. They are reached only through years of disciplined practice; they are themselves states generated by the brain, with their own neural signatures; and they replace one construction with another rather than delivering an unmediated view from nowhere. What they show is not that we can step outside the model, but that the model has parameters - including the self-world boundary - that are far more plastic than they feel. The cage has a door, but it opens onto another room, not onto the outside.

5. The Committee of Selves: Plurality and the Competition for Control:

The third illusion is the unity of the one who perceives. Introspection presents a single, continuous self - an author behind the thoughts, a chooser behind the choices, a witness behind the experiences. The science of the brain finds no such single author. It finds a coalition: many specialised processes running in parallel, often at cross purposes, with no central headquarters where "it all comes together." The felt unity is a product, not a starting point. In the user's own framing, there are *multiple personalities fighting inside the brain for control* - and the neuroscience, soberly stated, agrees.

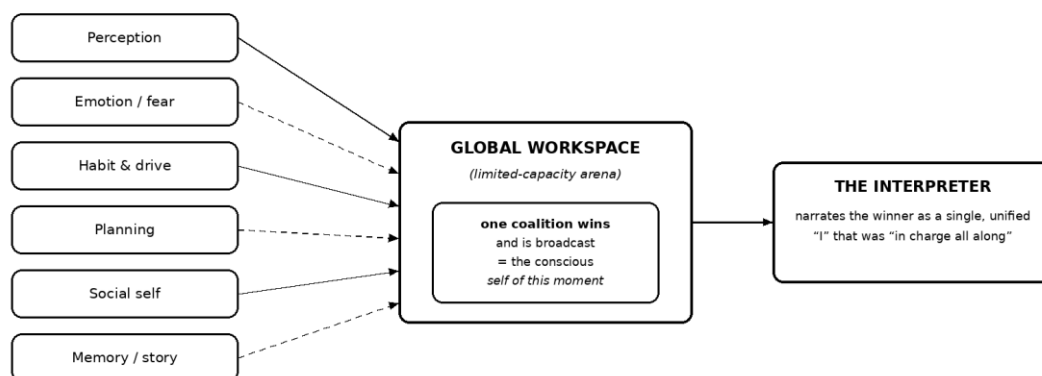
Minsky pictured the mind as a *society* of mindless sub-agents whose collective activity yields intelligence with no intelligent overseer at the centre (Minsky, 1986). Dennett pressed the point against the "Cartesian theatre" - the imagined inner stage where a unified self watches experience unfold - arguing instead for a *multiple-drafts* model in which many partial narratives are computed in parallel, with no privileged place or moment where they become "the" conscious account (Dennett, 1991). There is editing everywhere and a final cut nowhere.

Global-workspace theory gives the competition a mechanism. A large number of unconscious, parallel processors vie for access to a limited-capacity “workspace”; the coalition that wins is broadcast widely across the brain, and *that* broadcast is what we experience as the conscious content of the moment (Baars, 1988; Dehaene & Changeux, 2011). Consciousness, on this account, is not a controller but a stage with a single spotlight, and what occupies it is decided by competition rather than by a chooser. The self of any given instant is simply whichever coalition currently holds the floor. Figure 3 depicts the arrangement.

If unity is constructed, some process must be doing the constructing. The split-brain studies identify a culprit. When the two hemispheres are surgically disconnected, the left hemisphere - confronted with behaviour actually driven by the right, which it cannot access - does not report confusion; it instantly *invents* a plausible reason for the act and believes it. Gazzaniga calls this the *interpreter*: a left-hemisphere system that continuously weaves the outputs of countless independent processes into a single coherent story, the story we call our self (Gazzaniga, 2000, 2011). The unity of the self is the interpreter’s product, generated after the fact.

A single experiment makes the mechanism vivid. In one classic study a split-brain patient is shown a snow scene to the right hemisphere and a chicken’s foot to the left; asked to choose matching pictures, the left hand points to a snow shovel and the right hand to a chicken. When the speaking left hemisphere - which saw only the chicken - is asked to explain the shovel, it does not say “I do not know.” It answers, smoothly and with full conviction, that the shovel is for cleaning out the chicken shed (Gazzaniga, 2000). The reason is fabricated, yet it is delivered with the same sincerity as any ordinary self-explanation, and the patient believes it. The interpreter does not read off the true causes of behaviour; it manufactures plausible ones and presents the fabrication as self-knowledge. What such an experiment exposes in an unusual brain is, on this account, running constantly in the ordinary one.

Figure 3. The committee of selves: competition for control



There is no central controller. Many specialized processes compete; whichever coalition prevails is broadcast and momentarily becomes “the self.” A left-hemisphere interpreter then stitches the winners into one continuous story, producing the convincing - but constructed - sense of a single author behind our thoughts and acts.

Fig. 3. *The committee of selves. Many specialised processes compete for the global workspace; the winning coalition is broadcast and momentarily constitutes “the self,” while an interpreter narrates the succession of winners as a single, continuous author.*

Convergent findings show how routinely this post-hoc authorship occurs. People confidently explain choices they demonstrably did not make for the reasons they give, reporting more than they can know about their own minds (Nisbett & Wilson, 1977). Measurable brain activity predicts a “spontaneous” decision before the person is aware of having decided, suggesting that conscious will often ratifies an action already underway rather than initiating it (Libet, 1985). And everyday cognition runs on two loosely integrated systems - a fast, automatic one and a slow, effortful one - whose disagreements we experience as inner conflict (Kahneman, 2011). The self that takes credit for the verdict mostly arrives after the deliberation is over. It is, in Dennett’s phrase, a *center of narrative gravity*: not a thing in the brain, but a useful abstraction the narrative posits to organise itself (Dennett, 1992).

6. Synthesis: A User Illusion We Cannot Refuse:

The three claims compose a single picture. We inhabit a model of the world (Section 2); we cannot experience it as a model, because it is transparent (Section 4); and the self who seems to inhabit it is itself a construction, narrated into unity from a contest of parts (Section 5). Binding them is the keystone result (Section 3): the very boundary between that self and that world is a tunable neural construction, demonstrated by the fact that it can be dissolved (Saketh et al., 2026). Table 1 sets the everyday convictions against what the evidence indicates.

Table 1. *Three everyday convictions and what the evidence indicates.*

Everyday conviction	What the evidence indicates	Principal support
I perceive the world as it is.	I perceive a generative model constrained by, but not identical to, sensory input.	Predictive processing; controlled hallucination (Clark, 2013; Hohwy, 2013; Seth, 2021)
I stand outside my experience and observe reality directly.	The model is transparent; I cannot perceive it as a model, and knowing this does not change it.	Self-model theory; cognitive impenetrability (Metzinger, 2003; Gregory, 1997)
A single self perceives and decides.	A coalition of competing processes is narrated, after the fact, into one author.	Multiple drafts; global workspace; the interpreter (Dennett, 1991; Baars, 1988; Gazzaniga, 2011)
Self and world are fundamentally separate.	The boundary is a neural construction that dissolves when its timescale signature is equalised.	Neuro-temporal non-duality (Saketh et al., 2026)

The unifying notion is what some have called the *user illusion*: like the desktop on a computer, conscious experience is a simplified, manipulable interface that conceals the processes generating it (Hoffman, 2019; Metzinger, 2009). The illusion is not a flaw to be engineered away. It is what makes a brain usable to itself. A creature that had to perceive the wavelengths, the inferential machinery, and the parliament of sub-agents in their full complexity could not act in time to survive. The model *is* the simplification that makes a life possible. That it is false in detail is the price of its being useful at all.

This is why the empirical demonstration of non-duality matters so much to the argument. A construction that could never be altered would be indistinguishable, from the inside, from reality itself; the claim that experience is constructed would be unfalsifiable. But the self–world boundary *can* be altered, measurably and reproducibly, and altering the neural signature alters the experience in lockstep (Saketh et al., 2026). The exception proves the rule. We can briefly loosen the construction - and in doing so we confirm that it was a construction all along.

7. Objections and Boundaries:

Three objections deserve a direct answer. The first: **if perception is a lie, why does it work so well?** The answer is that “controlled hallucination” is controlled. The model is continuously disciplined by sensory error and selected over evolutionary time for fit between action and consequence (Friston, 2010; Hoffman, 2019). To say the model is constructed is not to say it is arbitrary, and certainly not to endorse solipsism: there is a world, and it constrains us at every step - we simply never meet it except through the model. Constructed and reliable-enough are wholly compatible.

The second: **is “lie” not too strong a word?** Strictly, yes. The brain does not deceive us with intent; it simplifies, predicts, and fills in, and the result is usually adaptive. “Lie” is a deliberate overstatement meant to puncture the default of naive realism - to mark the genuine and systematic gap between what we live and what is the case. A more temperate term is *necessary fiction*. But the moral force of the stronger word is earned: the model presents itself as the unvarnished truth, and that self-presentation is exactly the part that is false.

The third: **does explaining the construction explain consciousness itself?** It does not. Everything argued here concerns the *structure and contents* of experience - how the model is built, why it is opaque to itself, how the self is assembled. It leaves untouched the harder question of why any of this construction should be *felt* at all, why there is something it is like to be the system running the model (Chalmers, 1995). The present argument dissolves the illusion of direct access; it does not dissolve the mystery of experience. Those are different problems, and intellectual honesty requires keeping them apart.

8. Conclusion:

We are, each of us, sealed inside a model we did not choose and cannot fully see: a constructed world that presents itself as reality, a constructed self that presents itself as the author, and a felt separation between the two that presents itself as the bedrock of existence. The mind does not show us what is there; it shows us what is useful and hides the showing. This is not a defect of unusual or damaged minds. It is the design of the ordinary one.

The contribution of recent neuroscience is to move this claim from philosophy into measurement. By showing that the self–world boundary tracks the brain’s intrinsic timescales, and that equalising those timescales dissolves the boundary in both brain and experience, Saketh, Northoff, and colleagues have made the constructedness of the deepest layer of experience an empirical fact rather than a speculation (Saketh et al., 2026). The bars of the cage, it turns out, are made of neural time - and with great effort they can be bent. For most of us, most of the time, they will not be. But there is value simply in knowing they are bars: a measure of humility about our certainties, a lighter grip on the stories the interpreter tells, and

a more honest relationship with a reality we will never meet face to face, but only ever through the mind that, to serve us, must lie.

REFERENCES

- Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge University Press.
- Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.
- Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.
- Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200–227.
- Dennett, D. C. (1991). *Consciousness explained*. Little, Brown.
- Dennett, D. C. (1992). The self as a center of narrative gravity. In F. S. Kessel, P. M. Cole, & D. L. Johnson (Eds.), *Self and consciousness: Multiple perspectives* (pp. 103–115). Lawrence Erlbaum.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- Frith, C. D. (2007). *Making up the mind: How the brain creates our mental world*. Blackwell.
- Gazzaniga, M. S. (2000). Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition? *Brain*, 123(7), 1293–1326.
- Gazzaniga, M. S. (2011). *Who's in charge? Free will and the science of the brain*. Ecco/HarperCollins.
- Golesorkhi, M., Gomez-Pilar, J., Zilio, F., Berberian, N., Wolff, A., Yagoub, M. C. E., & Northoff, G. (2021). The brain and its time: Intrinsic neural timescales are key for input processing. *Communications Biology*, 4, 970.
- Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society B*, 352(1358), 1121–1127.
- Hoffman, D. D. (2019). *The case against reality: Why evolution hid the truth from our eyes*. W. W. Norton.
- Hohwy, J. (2013). *The predictive mind*. Oxford University Press.
- Josipovic, Z. (2019). Nondual awareness: Consciousness-as-such as non-representational reflexivity. *Progress in Brain Research*, 244, 273–298.
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8(4), 529–539.
- 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.
- Minsky, M. (1986). *The society of mind*. Simon & Schuster.

- Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*, 8(3), 102–107.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, 38, 433–447.
- Saketh, M., Sasidharan, A., Ventura, B., Venugopal, R., Bauer, C. C., Tewarie, P. K. B., Nagendra, R. P., Mehrotra, S., John, J. P., Subramaniam, B., Laureys, S., Kutty, B. M., & Northoff, G. (2026). Non-duality in brain and experience of advanced meditators - Key role for intrinsic neural timescales. *Communications Biology*. <https://doi.org/10.1038/s42003-026-10473-9>
- Seth, A. K. (2021). *Being you: A new science of consciousness*. Faber & Faber.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.
- von Helmholtz, H. (1925). *Treatise on physiological optics* (J. P. C. Southall, Trans.; Vol. 3). Optical Society of America. (Original work published 1867)