

## Unconsciousness and Self-Consciousness: Toward a Layered Account of Mind

Jitendra Chandolia\*

Research Scholar, Department of Philosophy, University of Rajasthan, Jaipur, Rajasthan, India.

\*Corresponding Author: jitendrachandolia18@gmail.com

### ABSTRACT

This paper develops a layered account of how unconscious mentality relates to self-consciousness. Historically, Leibniz and Kant already suggested that conscious thought rides on sub-personal processes and a formal “I.” Contemporary work refines this picture: higher-order and self-representational views tie consciousness to self-representation, while global-workspace models distinguish broadcast access from background computation. I argue, from debates on unconscious belief and qualia, that first-person authority is achieved through interpretation and that denying unfelt qualitative states risks epiphenomenalism. Evidence from blindsight and embodiment-body-ownership illusions, interoceptive prediction-shows minimal selfhood is scaffolded by largely unconscious processes. Dissociations between experience and meta-awareness caution against equating consciousness with report; a cautious pluralism follows.

**Keywords:** *Minimal Self, Meta-Consciousness, Unconscious Qualia, Embodiment, Blindsight.*

### Introduction

Philosophy, psychoanalysis, and cognitive science converge on a deceptively simple question: how does what the mind does **without us**-its background processing, habits, biases, and hidden motives-relate to the distinctive way experience is **for us**, namely self-consciously? Much turns on how we draw the line between the **unconscious** and **self-consciousness**. Do unconscious states have any built-in self-reference, or does self-consciousness arise only when states are made available to reflection and report? Does phenomenality itself-what it is like-ever occur without awareness? Contemporary debates answer these questions in sharply different ways, but a synoptic view suggests a layered picture in which unconscious processes are not merely the absence of consciousness, but part of the very scaffolding that makes self-consciousness possible (Crane 2017; Schaefer and North-off 2017; Kastrup 2017; Leite 2024; Coleman 2025).

The idea that mental life extends beyond the reach of awareness long predates Freud. Leibniz’s *petites perceptions* posited minute, unfelt representations that compose conscious life and sometimes rise to clarity by degrees; consciousness is thus continuous with a wider field of sub-personal activity (Bender 2020). On this picture, the border between the conscious and the nonconscious is not a wall but a gradient: innumerable faint perceptions lie below the threshold, summing into noticeable experience. Kant then sharpened the link between consciousness and the self with the transcendental unity of apperception-the formal “I think” that must be able to accompany representations-yet left open whether all consciousness *phenomenally* includes a felt “for-me-ness,” as opposed to being merely unified under an abstract “I” of judgment (Kant 1781/1787; see discussion in Tognazzini 2024). These two moves-Leibniz’s graded background and Kant’s formal self-continue to shape contemporary options: the mind may teem with sub-personal activity, while the conscious self may be present in more than one sense.

Twentieth-century currents divide and enrich this inheritance. Phenomenology emphasizes that ordinary consciousness is minimally self-present: experiences come marked by a pre-reflective “for-me-ness,” even when we are not explicitly thinking about ourselves (Zahavi 2005; Gallagher 2000). This “mineness” is not a judgment about myself but is woven into the way the world is given: the pain hurts **me**, the melody is heard **by me**. Psychoanalysis, by contrast, insists that thought, affect, and action are

massively shaped by processes **outside** introspection; desires and beliefs can be efficacious and yet unknown. For Freud the unconscious is structured and dynamic; for later theorists it becomes increasingly cognitive. Either way, introspection is no longer the arbiter of mentality (Kastrup 2017).

Current philosophy often cashes out the conscious–unconscious contrast in representational terms. **Higher-order** and **self-representational** theories claim that a mental state is conscious only when it is appropriately self-represented-by a higher-order thought about it or by self-presentation within the state itself. On this view, unconscious states simply lack the relevant self-representation; conscious states, by contrast, are reflexive in structure (Carruthers 2008; Kriegel 2009). Cognitive neuroscience adds a complementary lens: **global workspace** accounts treat consciousness as a matter of information being globally broadcast for control, report, and reasoning, leaving much sophisticated computation outside the “spotlight” (Dehaene 2014; Mashour et al. 2020). Ned Block’s influential separation of **phenomenal** from **access** consciousness warns us not to equate “what it’s like” with what is poised for report, a distinction that will matter when we test whether apparently unconscious capacities are in fact faint, unstable, or degraded forms of consciousness (Block 1995).

The case of **belief** shows why this terrain matters for self-knowledge. Eric Marcus has argued that belief is essentially self-conscious: to believe *p* is normally to *non-observationally* know that one believes *p*; more boldly, he extends this to **unconscious beliefs**, suggesting they carry an unconscious self-knowledge of their own presence (as discussed in Leite 2024). Adam Leite challenges both the empirical reach and the conceptual necessity of this view. Clinically, patients often *discover* an unconscious belief in therapy with genuine surprise; at other times they acknowledge a dim, inchoate familiarity- “I always sort of knew”-and the difference matters diagnostically and therapeutically (Leite 2024). Conceptually, the route from believing *p* to being able to honestly avow *p* is easily blocked by ambivalence, conflict, or context; so the inference from belief to built-in self-knowledge is too strong (Leite 2024). If Leite is right, the mind can host efficacious propositional attitudes that are not yet owned by the first person; self-consciousness arrives through interpretation and articulation, not as a standing property of every attitude.

Tim Crane offers a compatible explanation of why **unconscious belief** feels different from the conscious kind we avow. Attributions of unconscious belief, he argues, are modeling moves that pick out strands in a person’s **worldview**-a diffuse, holistic, and often indeterminate background orientation-rather than positing sentence-like states stored “below decks” (Crane 2017). When a person brings a belief to consciousness, they are articulating and stabilizing a portion of that worldview; the explicit “I believe that...” form is less a retrieval than a self-ascriptive act that imposes determinate content on a nebulous backdrop. On this picture, the self is present unconsciously as an *unarticulated point of view*, but explicitly self-referential belief is a product of conscious uptake (Crane 2017). The payoff is twofold: we can respect the efficacy of unconscious attitudes without over-intellectualizing the mind, and we can explain why introspection sometimes **confabulates**-not because it is defective, but because making sense of oneself is partly constructive.

If belief invites us to lower the bar for built-in self-knowledge, **qualia**-paradigmatically felt pain, color, or affect-tempt us to raise it. Many assume that qualitative character is, by definition, **conscious**. Sam Coleman argues that this assumption incurs a heavy cost. Consider patterns like **Restless Legs Syndrome**: when awake, a subject’s unpleasant creeping sensation prompts leg movement; during sleep, similar movements occur with no reported feeling. If the same neural basis produces similar downstream effects both with and without felt unpleasantness, either we accept that the conscious quality was **causally idle** when awake-an epiphenomenalist conclusion-or we allow that the quality may also occur **unconsciously** during sleep and do the same causal work (Coleman 2025). Coleman’s dilemma-**unconscious mental qualities or epiphenomenalism**-has bite beyond this case, given masked priming, affective influences without report, and imagery-driven behavior. To deny unconscious qualia one must either sever the tie between qualitative character and its usual neural basis or show that the sleep-case is not genuinely the same phenomenon. For many, positing **unfelt** qualitative states is the less damaging option (Coleman 2025). If so, phenomenality may come in grades and not always be accompanied by meta-awareness or recall.

The empirical poster child for the unconscious is **blindsight**: patients with primary visual cortex damage deny seeing yet correctly discriminate features in the blind field (Weiskrantz 1986; Ajina and Bridge 2016). At first pass, this looks like **perception without consciousness**. Yet recent work complicates the inference. Some argue blindsight involves **degraded** or atypical consciousness-hunches, feelings of presence, fleeting sensations-too weak or non-visual to count as ordinary seeing on coarse

self-report scales (Phillips 2020; Peters et al. 2017). The methodological lesson is twofold: robustly establishing *perception present* and *consciousness absent* is hard, and introspection can miss low-grade phenomenality. However that dispute falls out, blindsight reinforces the need to distinguish consciousness from access and to treat reportable self-consciousness as only one window onto experience. If blindsight is sometimes degraded consciousness, then some “unconscious” successes reflect **partial** awareness; if it is truly unconscious, we either accept a content-without-feel perception or revisit the neural–phenomenal mapping that makes such cases puzzling.

A complementary path to the unconscious–self relation proceeds through **embodiment**. Phenomenology has long distinguished a **minimal**, pre-reflective self-embodied “for-me-ness”-from the **narrative** or reflective self that evaluates, tells stories, and avows (Gallagher 2000; Zahavi 2005). Neuroscience gives this distinction traction. Schaefer and North-off associate conscious self-reference with **cortical midline structures** involved in autobiographical thought while emphasizing an **unconscious, embodied self** rooted in sensorimotor and interoceptive processing that silently scaffolds conceptual cognition-including moral cognition via bodily metaphors like *moral purity is physical cleanliness* (Schaefer and North-off 2017). Experimental illusions such as the **Rubber Hand Illusion** and clinical syndromes like **somatoparaphrenia** show that the sense of body-ownership-and with it, minimal selfhood-is labile, multisensory, and partly sub-personal (Botvinick and Cohen 1998; Feinberg and Venneri 2014). Thomas Metzinger’s self-model theory generalizes the point: the “self” is a transparent, predictive model; there is *no* self beyond these processes, many of which operate beneath introspection (Metzinger 2003; Blanke and Metzinger 2009; Limanowski and Metzinger 2013). Pathologies of depersonalization and derealization, now often framed in terms of interoceptive prediction, make vivid how shifts in bodily inference can estrange the subject from her own experience without any deliberate self-reflection (Gatus et al. 2022). In all of these cases, the **unconscious body** is not opposed to the self; it is a **constituent** of it.

A further wrinkle concerns the temporal and structural relation between **experience** and **meta-awareness**. Readers can traverse pages in a “zoned-out” mode, only later “waking up” to that fact; mind-wandering studies show repeated dissociations between having an experience and *noticing* that one is having it (Schooler 2002; Smallwood and Schooler 2015). The upshot is that **introspective reportability** is intermittent and fallible; conscious experience can outrun what we can readily avow. Bernardo Kastrup presses the philosophical moral: if we tacitly equate consciousness with what is **reportable** under meta-awareness, we will systematically **inflate** the domain of the “unconscious.” Many processes tagged “unconscious” may be **conscious but not meta-conscious**-co-conscious within the system but **inaccessible** to the introspective, reporting ego (Kastrup 2017). That does not settle every case; rather, it shifts the burden of proof. Before positing unconscious processing, we should ask whether our measures are sensitive to low-grade, unstable, or non-reportable phenomenality, and whether the relevant subject-think dissociation-might be another “center” within the same mind.

What, then, of **self-consciousness** proper? It is tempting to identify the self with whatever is available to reflective endorsement. But the foregoing materials suggest a more **layered** conception. At the base is **minimal selfhood**: a bodily, perspectival organization of experience defined by ownership, agency, and first-person perspective. This level is sustained by multisensory integration and interoceptive regulation that ordinarily run **outside** reflection. Above it stands a **narrative** or reflective self that interprets, explains, and avows; here we find Crane’s transition from worldview-level dispositions to articulated belief, and Leite’s insistence that first-person authority is an achievement rather than a default (Crane 2017; Leite 2024). Floating above both is **meta-consciousness**, which intermittently re-represents and polices experience, enabling explicit report and self-monitoring but also introducing blind spots (Schooler 2002). **Unconscious** processes populate and enable each layer differently. Bodily self-maintenance and sensorimotor predictions are constitutively **non-reflective** yet ground the minimal self; dispositional worldviews shape what can be consciously articulated as belief; even **phenomenality** may sometimes occur without meta-awareness (or else we accept epiphenomenalism about our most cherished feelings) (Coleman 2025; Kastrup 2017; Schaefer and North-off 2017).

This layered picture also clarifies several pressure points that drive contemporary disputes. First, the thesis that belief is **essentially** self-conscious overstates first-person authority; the mind contains efficacious attitudes that require **interpretation** before they become *mine* in the avowed, first-person way (Leite 2024; Crane 2017). Second, the insistence that **qualia** must always be felt risks epiphenomenalism and ignores evidence that qualitative influence can outrun report; allowing **unconscious mental qualities** preserves causal relevance without denying that felt experience often

matters in distinctive ways (Coleman 2025). Third, the allure of **unconscious perception** is tempered by demonstrations of **degraded** consciousness; but either outcome supports a graded, non-binary approach to awareness (Phillips 2020; Peters et al. 2017; Weiskrantz 1986). Finally, embodiment shows that self-consciousness is not just *about* the body; it is **made possible** by the body's largely unconscious dynamics (Schaefer and North-off 2017; Botvinick and Cohen 1998; Metzinger 2003).

If there is a unifying moral, it is a **cautious pluralism**. No single criterion-reportability, higher-order reflection, global broadcast-captures all that we intuitively mean by "being conscious," still less all that underwrites *self*-consciousness. Consciousness is not exhausted by what we can say; self-consciousness is not guaranteed by what we happen to feel. Instead, we should expect **graded phenomenality**, **interpretive** self-knowledge, and **embodied** scaffolding. Much of what we call the "unconscious" is not a ghostly underside but a family of ongoing processes-some dispositional and conceptual, some sensorimotor and interoceptive-that continuously shape the self we can avow. Self-consciousness is precious because it can stabilize, evaluate, and revise those processes; it is limited because it rides on them. The mind is **for us**-it presents a world in a first-person way-only because it is also, in important respects, **more than us**, a layered system whose lower levels we do not normally see. A philosophical account that honors both facts will be better equipped to explain why we sometimes fail to know our own minds, how reflective selfhood grows out of bodily and dispositional substrates, and why disputes about the unconscious so often reduce to disputes about what counts as *seeing*, *feeling*, or *knowing* in the first place (Crane 2017; Kastrup 2017; Schaefer and North-off 2017; Leite 2024; Coleman 2025).

## References

1. Ajina, Sara, and Holly Bridge. "Blindsight and Unconscious Vision: What They Teach Us About the Human Visual System." *Frontiers in Psychology*, vol. 7, 2016.
2. Bender, Sebastian. "Leibniz and the 'petites perceptions'." *Review of Metaphysics*, 2020.
3. Botvinick, Matthew, and Jonathan Cohen. "Rubber Hands 'Feel' Touch That Eyes See." *Nature*, vol. 391, 1998, p. 756.
4. Carruthers, Peter. "Higher-Order Theories of Consciousness." *The Stanford Encyclopedia of Philosophy*, 2008.
5. Coleman, Sam. "An Argument for Unconscious Mental Qualities." *Australasian Journal of Philosophy*, vol. 103, no. 1, 2025, pp. 216–234.
6. Dehaene, Stanislas. *Consciousness and the Brain*. Viking, 2014.
7. Feinberg, Todd E., and Annalena Venneri. "Somatoparaphrenia: Evolving Theories and Concepts." *Cortex*, vol. 61, 2014, pp. 74–80.
8. Gallagher, Shaun. "Philosophical Conceptions of the Self: Implications for Cognitive Science." *Trends in Cognitive Sciences*, vol. 4, no. 1, 2000, pp. 14–21.
9. Kriegel, Uriah. *Subjective Consciousness: A Self-Representational Theory*. Oxford UP, 2009.
10. Leite, Adam. "Some Challenges Raised by Unconscious Belief." *Philosophy and Phenomenological Research*, vol. 108, no. 3, 2024, pp. 838–843.
11. Limanowski, Jakub, and Thomas Metzinger. "Minimal Phenomenal Selfhood and the Free-Energy Principle." *Frontiers in Human Neuroscience*, vol. 7, 2013, article 547.
12. Metzinger, Thomas. *Being No One: The Self-Model Theory of Subjectivity*. MIT Press, 2003.
13. Peters, Maarten A. K., et al. "Does Unconscious Perception Really Exist? Continuing the ASSC20 Debate." *Neuroscience of Consciousness*, vol. 2017, no. 1, 2017.
14. Phillips, Ian. "Blindsight Is Qualitatively Degraded Conscious Vision." *Psychological Review*, vol. 127, no. 5, 2020, pp. 744–784.
15. Schaefer, Michael, and Georg North-off. "Who Am I: The Conscious and the Unconscious Self." *Frontiers in Human Neuroscience*, vol. 11, 2017, article 126.
16. Weiskrantz, Lawrence. *Blindsight: A Case Study and Implications*. Oxford UP, 1986.
17. Zahavi, Dan. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. MIT Press, 2005.

