How Can We Know Consciousness?

The epistemological question that comes before the hard problem

Elizabeth Carrio, Septembre 24, 2026

Why this question, and why now

I recently attended a conference on consciousness. Panel after panel, the discussions returned to the same oppositions: brain versus mind, science versus experience, materialism versus spirit. Strong positions were stated, personal experiences were offered as evidence, and much of the debate seemed to replay arguments that the literature settled, or at least reframed, decades ago.

What struck me was not the disagreement itself. Disagreement is healthy. It was the absence of a shared answer to a more basic question: what would count as knowing something about consciousness in the first place? Before we ask what consciousness is, or how it arises, we need to ask how it can become an object of knowledge at all. That is an epistemological question, and I want to argue that it is where much of the confusion in the field actually lives.

This question is also personal. My master's research used Organic Inquiry, a transpersonal method that treats the researcher's own experience as a legitimate source of data (Clements, 2004). I have therefore worked from inside a first-person methodology, with both its richness and its risks. This essay is an attempt to think clearly about both.

An old asymmetry

Consciousness is unusual among scientific objects. I have direct access to my own experience; you have access to it only through what I say, do, or display physiologically. Nagel (1974) gave this asymmetry its classic form: there is "something it is like" to be a conscious creature, and that subjective character does not obviously appear in any objective description. Levine (1983) named the resulting difficulty the *explanatory gap*: even a complete account of the neural events underlying pain seems to leave open why those events feel like anything at all. Chalmers (1995) later separated the "easy problems" of explaining cognitive functions from the "hard problem" of explaining experience itself.

These are not new ideas. That is precisely the point. Anyone entering the debate today inherits fifty years of careful work, and much of the circularity I observed comes from rediscovering these distinctions rather than building on them.

What a first-person report is, and is not

The asymmetry does not make subjective reports unscientific. It makes their status a methodological question. A report such as “I felt a presence” is information about experience, but it is also a behavioural event shaped by memory, language, attention, expectation, and the situation in which it is produced. Schwitzgebel (2008) has shown that people can be strikingly wrong about their own current experience, even about features they feel certain of.

Three responses to this problem have shaped the field. Dennett (1991) proposed *heterophenomenology*: take reports seriously as data about what subjects believe they experience, while remaining neutral about whether those beliefs are accurate. Varela (1996) proposed *neurophenomenology*: train subjects in disciplined first-person observation and use their refined descriptions to guide the analysis of neural data, so that each perspective constrains the other. A third line developed rigorous interview methods. Petitmengin (2006) designed a “second-person” interview technique that helps subjects describe the fine structure of experience without leading them, and Hurlburt and Schwitzgebel (2007) staged an extended debate between a proponent and a skeptic of experience sampling, which remains one of the most honest examinations of what introspection can and cannot deliver.

The lesson I draw is that first-person evidence is neither infallible nor worthless. Its value depends on how it is elicited, how it is checked, and how it relates to other forms of evidence. This is where Organic Inquiry, and transpersonal methods more generally, face their real test: not whether subjective experience matters, but whether the procedures used to gather and interpret it are transparent enough for others to evaluate.

The same caution applies to the absence of reports. Owen et al. (2006) showed that a patient diagnosed as vegetative could follow mental imagery instructions detectable by brain imaging. Behaviour, report, and consciousness can come apart. Their relationship is something to investigate, not something to assume.

Three levels of claims

Much confusion in consciousness debates comes from moving between different kinds of claims without noticing it. I find it useful to distinguish three levels:

> **Empirical observations**: a given neural pattern accompanies a reported experience.
>
> **Theoretical interpretations**: a proposed mechanism explains why that pattern relates to consciousness.
>
> **Ontological claims**: statements about what consciousness ultimately is.

A conclusion established at one level does not automatically transfer to another. Schurger and Graziano (2022) point out that many current theories *describe* the conditions associated with consciousness without *explaining* experience. That is a fair criticism of overconfident neuroscience. But the reverse move is equally unjustified: the fact that neuroscience has not explained experience does not demonstrate that consciousness is non-physical or independent of the brain. A gap in explanation is an invitation to further inquiry, not evidence for any particular alternative.

This symmetry matters to me because it cuts across the camps. It asks the same discipline of the materialist and of the spiritually inclined researcher.

When theories cannot be tested against each other

A survey of 166 consciousness researchers found deep disagreement about how to define consciousness, how to measure it, and which theories are most promising (Francken et al., 2022). Seth and Bayne (2022) identify four major families of theories: higher-order theories, global workspace theories, re-entry and predictive processing theories, and Integrated Information Theory. These families differ not only in their proposed mechanisms but in what they are trying to explain.

Two epistemological problems follow. The first is theory-ladenness: what counts as a valid measure of consciousness depends on the theory one already holds. The second is underdetermination: the same data can often be accommodated by several theories, especially when auxiliary assumptions can be adjusted after the fact. Negro (2024) argues, following Lakatos, that we should judge theories by the boldness of their predictions and the development of their research programmes over time, not merely by whether they can be made compatible with results. A theory that predicts is epistemically stronger than one that accommodates.

Adversarial collaboration: making disagreement productive

The COGITATE project offers one answer. Proponents of Integrated Information Theory and Global Neuronal Workspace Theory agreed in advance on conflicting predictions, which a theory-neutral team then tested with fMRI, MEG, and intracranial recordings in 256 participants (Cogitate Consortium et al., 2025). Neither theory won outright. Both saw some predictions confirmed and important claims challenged.

I find this outcome more interesting than a victory would have been. The value of the design lies in committing to predictions before seeing the data, which removes the easy escape of reinterpretation afterwards. The result is not agreement, but disagreement that is now constrained by evidence. That is what progress can look like in a field this difficult.

Some researchers argue that testing grand theories is not enough. Overgaard et al. (2025) propose complementing it with construct-based and data-driven approaches: breaking large theoretical notions into smaller measurable components, and letting rich datasets reveal patterns that no existing theory anticipated. None of these approaches is theory-free, but each creates another route by which evidence can challenge the way the field has organized its concepts.

What this means for transpersonal research

Transpersonal psychology has a particular stake in this discussion, because it studies precisely the experiences that strain ordinary methods: meditation, hypnosis, absorption, psychedelic states, mystical and exceptional experiences. It also has resources of its own that are too rarely brought into the conversation.

Tart (1972) argued in *Science* that altered states might require "state-specific sciences": communities of trained observers who can enter a given state, report on it, and check each other's observations, much as other sciences rely on trained perception. The proposal has well-known difficulties, but it was an early and serious attempt to give first-person inquiry intersubjective discipline. Ferrer (2002) later criticized transpersonal theory for treating inner experience as a private window onto a pre-given spiritual reality, and proposed a participatory view in which such events are co-created between person and world. Whatever one thinks of his conclusions, his critique names exactly the risk I observed: treating the felt authority of an experience as proof of its interpretation.

I see two symmetrical errors to avoid. The first is **reductionism by convenience**, where whatever can be measured objectively is assumed to be the whole phenomenon. The second is **phenomenological overreach**, where the intensity or conviction of an experience is taken as sufficient evidence for a metaphysical claim. Transpersonal research is most credible when it refuses both: when it takes experience seriously and still subjects its interpretations to shared scrutiny, triangulating reports with behavioural, physiological, and neural data wherever possible. Discrepancies between these sources are not failures. They are often where the most interesting findings lie.

Pluralism with discipline

Consciousness is probably too complex for a single method. But methodological pluralism is not relativism. Each method has specific inferential limits. First-person reports reach phenomenal character but are vulnerable to error and interpretation. Behavioural measures are observable but do not map uniquely onto experience. Neural data require theory to be interpreted as evidence of consciousness. Computational models formalize relationships but

inherit the assumptions built into their constructs. The task is not to merge these into one measure, but to establish how each constrains the others.

Seen this way, progress does not only mean arriving at the correct theory. It also means sharper definitions, better measures, experiments that can discriminate between rival views, and clearer awareness of what each kind of evidence can and cannot support.

The question I would bring to the next conference

Rather than asking once more whether consciousness is physical or not, I would propose a different question for the next panel:

> *What kinds of evidence can legitimately support which kinds of claims about consciousness, and how can different methods be combined without letting any one of them define the phenomenon in advance?*

This question does not solve the hard problem. It does something more modest and, I think, more useful: it clarifies the conditions under which a science of consciousness, including a transpersonal one, could make progress. Disagreement will not disappear. The aim is to make it more informative.

References

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200–219.

Clements, J. (2004). Organic inquiry: Toward research in partnership with spirit. *The Journal of Transpersonal Psychology, 36*(1), 26–49.

Cogitate Consortium, Ferrante, O., Gorska-Klimowska, U., Henin, S., Hirschhorn, R., Khalaf, A., … Melloni, L. (2025). Adversarial testing of global neuronal workspace and integrated information theories of consciousness. *Nature, 642*, 133–142. https://doi.org/10.1038/s41586-025-08888-1

Dennett, D. C. (1991). *Consciousness explained*. Little, Brown.

Ferrer, J. N. (2002). *Revisioning transpersonal theory: A participatory vision of human spirituality*. State University of New York Press.

Francken, J. C., Beerendonk, L., Molenaar, D., Fahrenfort, J. J., Kiverstein, J. D., Seth, A. K., & van Gaal, S. (2022). An academic survey on theoretical foundations, common assumptions and the current state of consciousness science. *Neuroscience of Consciousness, 2022*(1), niac011. https://doi.org/10.1093/nc/niac011

Hurlburt, R. T., & Schwitzgebel, E. (2007). *Describing inner experience? Proponent meets skeptic*. MIT Press.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly, 64*(4), 354–361. https://doi.org/10.1111/j.1468-0114.1983.tb00207.x

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 83*(4), 435–450. https://doi.org/10.2307/2183914

Negro, N. (2024). (Dis)confirming theories of consciousness and their predictions: Towards a Lakatosian consciousness science. *Neuroscience of Consciousness, 2024*(1), niae012. https://doi.org/10.1093/nc/niae012

Overgaard, M., Fazekas, P., Kirkeby-Hinrup, A., Sandberg, K., & Wiese, W. (2025). There can be more to consciousness research than theory testing. *Communications Psychology, 3*, Article 172. https://doi.org/10.1038/s44271-025-00357-9

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science, 313*(5792), 1402. https://doi.org/10.1126/science.1130197

Petitmengin, C. (2006). Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences, 5*(3–4), 229–269. https://doi.org/10.1007/s11097-006-9022-2

Schurger, A., & Graziano, M. S. A. (2022). Consciousness explained or described? *Neuroscience of Consciousness, 2022*(1), niac001. https://doi.org/10.1093/nc/niac001

Schwitzgebel, E. (2008). The unreliability of naive introspection. *The Philosophical Review, 117*(2), 245–273. https://doi.org/10.1215/00318108-2007-037

Seth, A. K., & Bayne, T. (2022). Theories of consciousness. *Nature Reviews Neuroscience, 23*, 439–452. https://doi.org/10.1038/s41583-022-00587-4

Tart, C. T. (1972). States of consciousness and state-specific sciences. *Science, 176*(4040), 1203–1210. https://doi.org/10.1126/science.176.4040.1203

Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies, 3*(4), 330–349.