

The Hard Problem Dissolved — But Into What?

A Critical Response to Carlo Rovelli’s

“There Is No ‘Hard Problem of Consciousness’”

Aernoud Dekker

May 2026

Version 1.0

Abstract

Carlo Rovelli’s recent essay in *Noema*¹ argues that the so-called hard problem of consciousness is a confusion produced by treating science as a view from outside the world rather than as a perspectival account constructed from within it. He is largely right about what he is against: Cartesian substance dualism, the philosophical-zombie thought experiment as a proof of irreducibility, and the residual nostalgia for a transcendent soul. Yet his positive move — that mental properties are obtained by *subtraction* from a complete physical account — stops short of the position his own premises imply. This response, written from within the Ignorant Observer Framework (IOF) and its non-dual interpretive layer, grants Rovelli the ground he shares with IOF and then presses on six places where his deflation goes too far. The aim is not to restore the hard problem in Chalmers’s form, but to show that what was right in the hard problem — the asymmetry between the describing subject and any describable object — survives Rovelli’s argument intact, and indicates a non-dual reading rather than a flat physicalist monism.

Reading Rule

This is primarily a philosophical response. Its empirical claim is conditional and enters only in Section 7: *if* the BLQC visibility prediction

$$V_{\text{obs}} = V_{\text{std}} e^{-\sigma_{\theta}^2/2}, \quad t_{\text{break}} \propto \frac{1}{h_{\text{KS}} - C_{\text{eff}} \ln 2},$$

is experimentally confirmed, then the philosophical pressure of Section 7 becomes a structural pressure rather than a metaphysical preference. Sections 2–6 do not depend on that result.

“Any account is perspectival because knowledge is always embodied.”

— Carlo Rovelli

“And what is embodied cannot, from within the body, give an account of the embodying.”

— the reply IOF offers

¹<https://www.noemamag.com/there-is-no-hard-problem-of-consciousness/>

1 What Rovelli Gets Right

Before pressing, the agreement is worth marking. Rovelli’s essay is not a defence of strong physicalism in the Dennett mode; it is a deflation of a particular dualistic move, and on that ground IOF shares almost all of his terrain.

Both views reject Cartesian substance dualism. Both reject the notion that consciousness is an *extra ingredient* added to a complete physical account. Both treat science as a situated, embodied activity rather than a view from nowhere. Rovelli’s line that “any knowledge is perspectival” and that “the world I access is the information I have about it” is structurally identical to IOF’s claim that quantum mechanics is operationally complete as epistemology for a bounded observer with finite effective capacity C_{eff} . His insistence that the dualism between first-person and third-person description is “a normal perspectival difference” parallels IOF’s reading (§5 of the main paper) that what we call *collapse* is an epistemic update inside an observer, not a physical event in the world. Even his negative claim about the explanatory gap — that it is the wrong shape of question — is one IOF endorses, though for different reasons.

So the disagreement that follows is not the disagreement of a dualist with a monist. It is the disagreement of a non-dual monist with a physicalist monist who has not noticed how much of physicalism his own argument has already abandoned.

2 Two Perspectives Require a Bridge Principle

Rovelli’s central deflationary move is this: “The dualism between a first-person description of experience and a third-person (or scientific) account of the same is a normal perspectival difference: the same brain phenomenon as experienced by that same brain itself, or by another.”

The sentence is doing more work than it appears. To say “the same brain phenomenon, seen one way from within and another way from without” is to identify a numerically single object across two views. That identification is not free; it requires some account of what licenses calling the two views views of one thing. Rovelli can reasonably reply that no view from nowhere is needed for this — only an embodied scientific practice that links perspectives by stable invariants (the same neurons, the same dynamics, the same coarse-grained patterns appearing under different couplings). That is a fair reply, and the section’s claim should be stated accordingly: Rovelli owes us a bridge principle that explains how sameness across perspectives is licensed without quietly reinstating the privileged physical description as the standard against which the perspectives are compared.

The risk is real. The candidate bridge principles tend to fall into two families. Either the bridge is a third-person description treated as the canonical referent (in which case the first-person view is mapped onto it and the priority of the physical re-enters by the back door); or the bridge is whatever stable invariants the two perspectives happen to share, in which case the question of what makes them *the same thing* dissolves into “they co-vary lawfully,” which is less than the strong identity claim Rovelli’s argument needs to dispel dualism.

IOF, in the non-dual reading worked out in §17 of the main paper, supplies a different kind of bridge. The asymmetry between observer and observed is not a perspectival difference *within* a shared frame; it is the condition for there being a frame at all. The Vedāntic distinction between

turīya (the substrate as seen from the standpoint of experience) and *brahman* (reality itself, which makes any standpoint possible) names what Rovelli’s two-perspectives formulation needs and does not provide: an account of how the standpoint and the described co-arise, without granting either ontological priority. Physics can describe the boundary-facing aspect of the substrate. It cannot reach the reality that makes description possible, because any reaching is itself description — and that limitation is structural rather than a temporary gap in current science.

3 The Subtraction Account Inverts the Dependency

Rovelli is explicit that mental properties are not obtained by adding to a physical account but by subtracting from it. The complete account is physical; mental categories are coarse-grained physical categories. He writes that mental and spiritual entities “are not obtained by *addition* to a physical state, but by *subtraction* from a complete physical account.”

The order of priority matters. To make “subtraction” the right verb, the complete physical account has to come first ontologically. But on Rovelli’s own perspectivism, every account is produced *by* an observer of finite reach, *about* the appearance that observer has access to. The “complete physical account” is not a thing the universe possesses and we read off; it is a regulative ideal of a third-person description, and one no embedded observer has ever produced or, IOF would argue, can produce.

IOF makes this argument precise. The observer is a physical system with finite effective capacity C_{eff} and nonzero internal entropy-rate h_{KS} . The Data-Rate Theorem then puts a hard bound on what can be tracked — including by the observer about itself. The deterministic substrate ($|\psi\rangle, \{\xi_i\}$) is real in the sense that it figures in the formalism, but it is structurally inaccessible from within. What we have is always an epistemic theory shaped by the observer’s bounded resources. Quantum mechanics is then “operationally complete as epistemology while incomplete as ontology” — and this is not a coyly anti-realist hedge but a description of a real structural fact about embedded knowers.

If this is right, then Rovelli’s subtraction is the wrong direction. The mental is not a coarse-graining of a third-person complete description; the third-person description is a structured projection inside the field that the mental, in his vocabulary, occupies. To call the soul a subtraction from the physical assumes precisely what cannot be earned from a perspectival starting point: that the physical is the prior, complete category.

4 The Zombie Argument After IOF

Rovelli’s reply to Chalmers’s philosophical-zombie thought experiment is clever and partly right. He notes that a physically identical zombie would, by hypothesis, report consciousness, write essays about consciousness, and be convinced by introspection of its own experience exactly as we are. Therefore my own introspective certainty cannot, by itself, distinguish me from a zombie, and the argument from introspection “is self-defeating.”

Against a Chalmers who treats qualia as an extra ingredient *added on top* of a physical de-

scription, this works. But it does not work against IOF’s reading, which is a third option Rovelli does not consider. On the IOF reading, the zombie’s report would be structurally identical because what subjectivity *operationally presents as* at the vyāvahārika (empirical) level — i.e., what the empirical *jīva* or bounded observer is in this framework — includes a self-ignorance structure: the system cannot, from within, trace why its own basis $\theta(t)$ took the value it did. This is a claim about the operational structure of being-an-observer, not a definition of consciousness itself; the main paper is explicit that consciousness as such is not modelled by IOF, and that not every finite controller is therefore a subject. What IOF does model is the self-ignorance signature that any bounded knower carries. Both you and your hypothetical twin instantiate that signature. There is no further empirical fact that decides which of you is “really” conscious — because the question presupposes a comparison from a standpoint neither of you has.

What this means for Chalmers’s argument is that it does not establish ontological dualism, but it does establish something Rovelli ignores: that subjectivity is not exhausted by the third-person description of the brain, because that description is itself produced by a subject standing under the same self-ignorance constraint. The zombie scenario fails as proof of dualism and succeeds as a pointer to a structural asymmetry that flat physicalism elides.

5 Past Gap-Closures Don’t Buy This Closure

Rovelli’s most rhetorically effective move is the historical induction: heaven and Earth turned out to share one mechanics; living things and inert matter turned out to share one chemistry; humans and other animals turned out to share one lineage. The pattern is unmistakable: apparent metaphysical gaps repeatedly close. Therefore the consciousness gap will close too.

The induction equivocates between two kinds of gap. Call them type-A and type-B. A type-A gap is a discontinuity *inside* the third-person frame: two regions of the objective world initially described by different theories that turn out to fall under one. Heliocentrism, evolution, biochemistry, neuroscience-of-perception — each of these is a homogenisation *within* third-person description. A type-B gap is the relation between the third-person frame and the standpoint from which any frame is constructed. That is a different category. Every successful type-A closure in the history of science has presupposed a knower for whom the unified description is a description; none of them has addressed, let alone closed, the relation between that knower and the description.

The induction therefore has no purchase on the question. You cannot argue from “every gap inside the picture has closed” to “the gap between the picture and the picturer will close.” The latter is not the same kind of fact as the former, and treating it as if it were is a category error of the kind §20 of the main paper analyses through the rope-snake parable: the entire debate about how to describe the snake (one snake? many?) is conducted on the wrong side of the lamp.

IOF’s framework agrees that the type-A gaps will continue to close; that is normal science. It denies that this empirical track record gives *any* reason to expect type-B closure — because type-B closure would require an account that includes its own grounding, which Gödelian self-reference and IOF’s self-tracking bounds independently suggest is structurally unavailable to an embedded observer.

6 The Direction of the Epistemic–Ontic Circle

Rovelli writes: “We do not need to assume that the circle between epistemology (how we get knowledge) and ontology (what exists) requires a starting point. There is nothing wrong with its circularity: the world I access is the information I have about it, and I am part of that world.”

The circularity is real and IOF affirms it. But the circle is not benign in the way Rovelli suggests, because it has a direction. Any attempt to close the loop by objectifying the standpoint — by giving an account of the account-giver — produces a further appearance inside the very structure being described. This is the IOF version of an old Vedāntic point: the ego attempting to observe itself is structurally like the eye attempting to see itself. It produces another image, not direct knowledge.

IOF supplies a concrete physical instance of this open loop — not a general proof that no account-giver can be exhaustively accounted for, but one well-defined model in which the loop is provably unclosable. The model treats the observer’s basis-tracking control problem as implemented in the same finite physical substrate it is trying to track. The deficit

$$\kappa = h_{\text{KS}} - C_{\text{eff}} \ln 2$$

quantifies a specific tracking failure: when $\kappa > 0$ (chaos-wins), the observer cannot maintain a stable representation of its own effective measurement basis; when $\kappa < 0$ (capacity-wins), local tracking succeeds but the upstream trace of why the basis evolved as it did remains hidden. This is one mechanism by which a bounded knower fails to objectify itself; it is not the only conceivable one, and the general philosophical claim that self-reference is structurally productive of further appearance does not stand or fall with this specific dynamical model. What the model contributes is a concrete physical handle on a structure usually only argued for in the abstract.

Rovelli treats the circle as harmlessly self-supporting. IOF treats it as productive but essentially open: the closure is exactly what physics cannot accomplish, and acknowledging this is not a confession of dualism but of the structure of self-reference. What lies on the unclosable side is not a Cartesian *res cogitans*; it is the simple fact that any frame requires a frame-bearer the frame does not describe.

7 Two-Stage Empirical Pressure: BLQC and the Penrose Convergence

Sections 2–6 are philosophical. This one identifies where the disagreement may take empirical shape, in two stages whose status is deliberately distinguished.

Stage one: BLQC. IOF’s central testable prediction is that observable visibility under finite-rate basis tracking takes the form $V_{\text{obs}} = V_{\text{std}} e^{-\sigma_{\theta}^2/2}$, with characteristic breakdown timescale

$$t_{\text{break}} \propto \frac{1}{\kappa}, \quad \kappa = h_{\text{KS}} - C_{\text{eff}} \ln 2,$$

in the chaos-wins regime ($\kappa > 0$), up to threshold factors, with C_{eff} and h_{KS} independently variable in a controlled phase-reference experiment. This is the free-standing falsifiable claim. If BLQC confirms the bandwidth dependence on κ , then the structure of finite self-knowledge is doing physical work at a measurable scale — not as a metaphor for the asymmetry between knower and known, but as one operational realisation of it. That would not refute Rovelli’s view, but it would set a constraint his current framework has no role for: a regime in which the observer’s tracking deficit, rather than a coarse-graining choice, controls when superposition becomes operationally indistinguishable from classical randomness. If BLQC does not confirm the dependence, this empirical pressure simply does not arise, and Sections 2–6 stand as a purely philosophical reading.

Stage two: the Penrose convergence, conditionally upgraded. Until BLQC delivers, the numerical overlap between IOF’s self-knowledge timescale τ_{oss} and Penrose’s gravitational-collapse timescale τ_{OR} — both falling in the same mesoscopic window of $\sim 50\text{--}70$ ms at masses of order 10^{-15} kg (§§12, 22 of the main paper) — is an illustrative coincidence, no more. Two independently motivated timescales happen to land at the same scale; this is suggestive but not load-bearing.

If BLQC confirms the predicted dependence on κ , the overlap *changes category*. It is no longer an illustrative coincidence; it becomes a candidate structural convergence between an empirically measured epistemic limit and an independently proposed ontological instability scale. That is a different kind of claim, and a more demanding one for a deflationary reading to absorb. The convergence remains conditional on Penrose’s OR hypothesis (or a near-relative of it — the Diósi–Penrose family is itself empirically contested, and what matters here is the existence of some gravitationally motivated instability scale at this order, not the exact functional form). But the conditional has a different texture once stage one is settled: it is no longer “if you happen to take both pieces of speculation seriously,” but “given an empirically confirmed epistemic boundary, here is a quantitative ontological boundary that lands at the same scale.”

A deflationary view of the quantum-classical transition would then have to treat the convergence of two independently motivated mechanisms at one scale as accidental. That is a price; it is not a refutation. But the staging makes clear where the price is paid and what would have to happen for the bill to come due.

The asymmetry of the two stages. Stage one is falsifiable on its own and defensible whether or not stage two ever activates. Stage two is conditional, but the conditioning runs through stage one in a particular way: the philosophical move IOF wants to make — that the relation between epistemic limitation and ontic manifestation is structural rather than merely perspectival — gets a quantitative lever exactly when the staging completes. Rovelli’s framework, in its current form, has no analogue prediction at either stage. The absence is worth noting at stage one and is most striking at stage two, where the deflationary reading would have to absorb a structural convergence its premises do not anticipate.

8 Conclusion: Same Direction, Different Distance

Rovelli ends his essay with a generous line: “We have souls. We have an inner self. We can treat ourselves as transcendental subjects in the Kantian sense.” IOF would echo this without qualification. The disagreement is not about whether we have an inner life, but about what that fact says about the structure of reality.

Rovelli’s argument moves correctly in three steps and then stops. Step one: reject Cartesian dualism. Step two: refuse the view from nowhere. Step three: insist that knowledge is embodied and perspectival. These three steps are also IOF’s. But there is a fourth step his argument’s premises imply and his conclusion declines to take: if any account is from within, if every closure of the epistemic-ontic circle produces a further appearance, and if the standpoint from which description is made cannot itself be exhaustively described, then the right monism is not the flat physicalist one — “soul of the same basic nature as a thunderstorm” — but the non-dual one in which the soul and the thunderstorm are both appearances within a reality whose nature is consciousness rather than matter, where “matter” itself is consciousness’s structured self-appearance through avidyā.

This is a stronger monism than Rovelli’s, not a retreat to dualism. It does not multiply substances. It refuses, instead, to grant priority to the third-person description at exactly the point Rovelli’s perspectivism should refuse it. And it stages a falsifiable claim Rovelli’s framework does not: that finite-rate basis tracking governs when quantum coherence becomes operationally indistinguishable from classical randomness, and — conditional on that confirmation — that the boundary at which the observer loses self-knowledge of its basis coincides with a candidate ontological instability scale at the same mesoscopic regime. Stage one is testable on its own terms. Stage two becomes a structural claim once stage one is settled.

The hard problem, in Chalmers’s substance-dualist form, is dissolved. Rovelli is right about that. But the asymmetry the hard problem was pointing toward — the one that distinguishes a describing standpoint from the described — survives the dissolution. What IOF offers is a way to honour that survival without postulating a second substance: a non-dual reading in which the asymmetry is the structure of self-reference itself, made physical by the finite capacity of any embedded knower, and tested empirically by the BLQC programme.

We share more of Rovelli’s ground than the essay’s title suggests. We differ on how far his own argument should be allowed to travel.

“The eye that mistakes itself for the world cannot describe the eye.”