

What Survives When The Hard Problem Goes  
A Response to Carlo Rovelli —  
the Noema-voice version of *The Hard Problem Dissolved — But Into What?*

Aernoud Dekker

May 2026

Version 1.0

**Abstract**

This is the popular, magazine-register companion to the technical response *The Hard Problem Dissolved — But Into What?* It rewrites the same argument for a general audience in the voice of *Noema*. The thesis is unchanged: Carlo Rovelli’s recent essay is mostly right, and stops one step short of the conclusion its own premises imply. The hard problem in its Cartesian form is dissolved by his argument — but the asymmetry the hard problem was pointing toward, between the standpoint that describes and the world that is described, survives the dissolution. The right response is a stronger monism than Rovelli’s, not a retreat to dualism. And one piece of it is empirically testable.

*“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 offered here

Carlo Rovelli wants to settle a long-running argument. The mystery of consciousness, he says, is a confusion. There is no special difficulty in explaining how a brain becomes a mind, and certainly no need to invoke a soul made of finer stuff. The so-called “hard problem of consciousness” is, in his telling, a piece of intellectual furniture left over from the Middle Ages — a worldview in which mind and matter belong to different domains, and which keeps reasserting itself even after every reason for believing in it has been retired.

He is mostly right.

But Rovelli’s argument has a shape that interests me more than its conclusion. He doesn’t, in the end, win his case by appealing to brain science or to the explanatory power of physics. He wins it by undoing a way of looking. The hard problem, he says, depends on imagining science as a view from outside the world, against which our first-person experience cannot quite be lined up. Drop that picture — recognize that any account of reality is constructed from within it, by an embodied knower — and the problem dissolves. There is no view from nowhere; the gap between objective description and subjective experience is just the perspectival difference between watching a brain from outside and being one from inside.

I want to follow Rovelli’s argument exactly that far, agreeing with him at each step, and then ask what happens if we take it one step further than he does.

## Why It Matters

The hard problem is not just a parlour game. It sits underneath some of the most consequential questions of our moment. Whether an AI system can be conscious. Whether a fetus, an octopus, or a person in a deep coma is having an experience. Whether there is anything to fear, or to lose, in death. The way we answer the metaphysical question shapes the way we answer the ethical ones.

Rovelli’s deflation is one cultural response to this pressure: a confident insistence that the question has no special depth, that consciousness is “of the same basic nature as a thunderstorm.” His readers feel either relieved by this or quietly resistant to it. The relief is sensible. The resistance, I’ll argue, is also pointing at something real — not at a soul, but at a structural fact about what it means to give an account of anything from within the system being accounted for.

## Where We Agree

Three of Rovelli’s premises are also mine.

First, there is no Cartesian soul. The idea that the human person is composed of a physical body and a separate non-physical substance is wrong, and centuries of cumulative evidence — from neurology to evolutionary biology — have made this almost too obvious to state. If we want a place for the soul in our metaphysics, we will have to find it inside the world, not outside it.

Second, there is no view from nowhere. Any description of reality is produced by some particular observer, embedded in the reality they are describing, using cognitive tools shaped by the same processes that produced the rest of the world. Rovelli writes that “any knowledge is perspectival,” and any embedded knower will say the same.

Third, the gap between first-person experience and third-person description is not what dualists make it out to be. There is no second, ghostly substance that the brain happens to be made of in addition to its physical matter. Whatever consciousness is, it is not an extra ingredient added on top of a sufficient physical account.

So far, Rovelli and I are walking together. The disagreement starts where he says: “and so the so-called hard problem dissolves.” I want to ask what, exactly, dissolves with it.

## Bridge Trouble

Rovelli’s central move is to recast the first-person/third-person divide as a perspectival difference: “the same brain phenomenon” seen one way from inside, another way from outside. The sentence sounds harmless. It is doing more work than it admits.

To say “the same brain phenomenon” is to identify a numerically single object across two viewpoints. That identification is not free. Someone, somewhere, has to license the claim that the felt experience and the measurable neural pattern are the same thing rather than two correlated things. Rovelli’s strategy implies that no observer-from-outside is needed for this — only the everyday scientific practice of linking perspectives by what stays invariant across them: the same neurons, the same activity pattern, the same coarse-grained dynamics under different couplings.

That is a fair reply. The trouble is that the available bridges all have problems. Either the bridge is the third-person description treated as the canonical referent — in which case the first-person view is being mapped *onto* the physical, and the priority of the physical has been reinstated by the back door. Or the bridge is whatever invariants the two perspectives happen to share, in which case “they are the same thing” weakens into “they reliably co-vary” — which is much less than the strong identity claim Rovelli’s argument needs in order to settle anything.

This isn’t a knockdown. It is the observation that Rovelli owes us a bridge principle, and the candidates either give back the dualism he is trying to dissolve, or give back something so weak it dissolves the position he is trying to defend.

## The Subtraction Problem

The same trouble appears more cleanly in another of Rovelli’s phrases. Mental properties, he writes, “are not obtained by *addition* to a physical state, but by *subtraction* from a complete physical account.” The mental is just the physical, described with fewer details. This is meant to deflate the temptation to imagine a separate mental substance. It does that. It also assumes something it never proves.

To make “subtraction” the right verb, the complete physical account has to exist first. It has to be available, at least in principle, as the prior reality from which mental categories are then derived. But on Rovelli’s own perspectivism, every account is produced *by* a finite observer, *about* the slice of reality 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, a horizon that no embedded knower has ever reached or, I want to argue, can reach.

The argument that no observer can reach it is not philosophical hand-waving. It is a consequence of treating observers seriously as physical systems. Any system that is going to track and describe the world has some finite information-processing bandwidth — a maximum rate, in bits per second, at which it can absorb, integrate and represent new states. And any system with internal dynamics has some characteristic rate at which its own state becomes unpredictable, even to itself. A control theorist will recognize the resulting deficit as familiar: when the rate of internal unpredictability outruns the bandwidth available to track it, the system cannot stably represent its own state. It cannot know, from within, why it is in the configuration it happens to be in.

This is not yet a theory of consciousness. It is a theory of the structural limit any finite knower meets when trying to know itself. And it shows that the “complete physical account” Rovelli’s subtraction requires is, in a quantifiable sense, structurally unavailable from inside. The third-person description we actually have is always an epistemic projection by an observer who cannot fully objectify themselves. To call the mental a subtraction from the physical is to assume that the physical is the prior, complete category. It is not. The physical, in the sense of the fully described, is itself a projection inside the perspectival field Rovelli has just told us we cannot escape.

## Zombies, Reread

Rovelli also offers an admirably clean reply to the philosophical-zombie thought experiment. David Chalmers asked us to imagine a being physically identical to a conscious person but with “nobody home” — no inner experience, no qualia. If such a being is conceivable, Chalmers argued, then consciousness must be something extra, irreducible to physics. Rovelli notes that a physically identical zombie would, by hypothesis, claim consciousness, write essays about it, and convince itself through introspection that it has inner experience exactly as we do. So introspection cannot, by itself, tell us we are not zombies — and the introspective argument for irreducible qualia is, he writes, “self-defeating.”

Against Chalmers, this works. Against a different position, it does not.

Suppose we take seriously the structural fact noted above: that any embedded observer carries a built-in self-ignorance — an inability to fully track from within why its own state took the value it did. On this reading, what we call subjectivity, *at the level of the empirical world we know from inside*, is not a separate substance riding on top of the physical. It is the way the world looks from within a finite knower whose self-tracking is incomplete. The zombie’s report would be structurally identical to ours not because consciousness is reducible to behaviour but because what we are pointing at, when we point at “having an inner life,” includes this self-ignorance signature — and both we and our hypothetical twin instantiate it.

This is a careful claim, and worth stating slowly. It does not say that consciousness *is* finite-bandwidth self-ignorance. The framework I’m working in is explicit that it does not model consciousness as such, and that not every limited information-processing system is automatically a subject. What it does model is the structural asymmetry between the standpoint that gives an account and the world that gets accounted for. That asymmetry is real, it is empirical, and it is what the zombie scenario was pointing at all along. Chalmers misnamed it as a substance. Rovelli

mistook its mention for a confession of dualism. It is neither.

## Two Kinds Of Gap

The most rhetorically effective move in Rovelli’s essay is the historical induction. Once we thought the heavens were made of different stuff than the Earth. Then Newton showed they obey the same mechanics. Once we thought living matter contained a special spark; biochemistry revealed it doesn’t. Once we thought humans were a separate creation; Darwin made us cousins of donkeys. The pattern is unmistakable: apparent metaphysical gaps repeatedly close. Therefore, Rovelli argues, the consciousness gap will close too.

I want to suggest that this induction quietly equivocates.

There are two kinds of gap in the story. The first kind — call it Type A — is a discontinuity *inside* the third-person picture. Two regions of the objective world initially described by different theories turn out to fall under one. Heliocentrism, evolution, biochemistry, neuroscience — every closure Rovelli cites is a unification *within* the third-person frame, the picture we paint of the world. The second kind of gap — call it Type B — is the relation between that picture and the standpoint from which the picture is painted. Closing a Type A gap is normal science. Closing a Type B gap would mean producing an account that included its own grounding: a description that made fully explicit the position of the describer. That is a different kind of task, and nothing in the history of Type A closures gives us inductive purchase on it.

This is not a mystical point. It is the same structural fact that Kurt Gödel turned into a theorem about formal systems and that contemplative traditions have been turning into liberation practices for thousands of years: a sufficiently expressive system cannot give a complete account of itself from within. Rovelli’s induction is unimpeachable for Type A gaps. It has no traction on Type B.

## A Circle With Direction

Late in the essay, Rovelli offers what is meant to be a reassuring image. The relationship between what we know (epistemology) and what exists (ontology) is, he says, a circle, and that’s fine. “The world I access is the information I have about it, and I am part of that world.” No starting point is needed.

The circularity is real and I accept it. But the circle is not benign in the way the image suggests, because it has a direction. Every attempt to close the loop by giving an account of the account-giver produces another appearance inside the very structure being described. This is the old Vedāntic point — that the eye attempting to see itself produces another image, not direct sight — but it is also a perfectly secular fact about self-reference.

In the framework I’ve been developing, this fact takes a concrete physical form. The observer’s basis-tracking control loop is implemented in the same finite physical substrate that the loop is trying to track. When the rate at which the substrate becomes unpredictable exceeds the bandwidth available for tracking, the system cannot stabilize a representation of its own basis. The deficit can be written compactly as a number, called  $\kappa$  — the gap between internal unpredictability and

tracking bandwidth — and the system’s failure to know its own state is governed by it. This is not the only way a bounded knower can fail to objectify itself, and the general philosophical point does not depend on 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 epistemic–ontic circle as self-supporting. I would say it is productive but essentially open. The closure he gestures at is what physics cannot, in principle, accomplish. Acknowledging this is not a confession of dualism. It is a recognition that any frame requires a frame-bearer the frame does not describe.

## A Testable Move

So far the disagreement has been philosophical. Now I want to point at the place where, if the framework I’ve been describing is right, it becomes empirical.

The same structural argument that locates the unclosable side of the epistemic circle predicts a quantitative effect in a laboratory. If the visibility of quantum coherence in mesoscopic interferometry experiments — the kind of experiment that asks whether a sufficiently large physical system can be in a superposition of two states — depends on how well the experimental apparatus can track its own measurement basis, then we can write down a prediction. Specifically: in the regime where the apparatus’s internal unpredictability exceeds its tracking bandwidth, visibility should break down with a characteristic timescale proportional to  $1/\kappa$ . The bandwidth and the unpredictability can be independently varied in a well-designed control experiment. This is a falsifiable claim, free-standing of any metaphysical interpretation.

If this prediction holds, something striking happens to a separate piece of physics. In the 1990s, Roger Penrose proposed that quantum superpositions might collapse spontaneously under their own gravitational weight — a hypothesis sometimes called objective reduction. The collapse timescale he predicted for a mesoscopic object of about  $10^{-15}$  kilograms falls in the range of 50 to 70 milliseconds. The timescale at which a low-bandwidth embedded observer or controller would lose track of its own measurement basis happens to fall in the same window.

Until the bandwidth-dependent visibility experiment is done, this overlap is a striking but unsupported coincidence. Two independently motivated timescales happen to land at the same scale. That’s suggestive and nothing more. But if the bandwidth experiment confirms the predicted dependence, the overlap changes category. It becomes a candidate structural correspondence between an empirically measured epistemic limit (the scale at which observers cannot objectify their own state) and an independently proposed ontological one (the scale at which superpositions become gravitationally unstable). A view of consciousness that treats first-person and third-person as smoothly interconvertible perspectives would have to absorb that convergence as accidental. That is a price; it is not a refutation, but it is a price.

Andrea Vinante’s collaboration and several other groups have been running ultra-cold cantilever experiments that bound the Penrose-style collapse hypothesis. The bandwidth-dependent visibility experiments are technically demanding but within reach of trapped-ion and superconducting platforms now in operation. The empirical question is concrete and falsifiable — not philosophical

speculation.

## The Step He Doesn't Take

What survives, then, when the hard problem in its Cartesian form dissolves?

Not a second substance. Not a ghost. Not a soul carved out of finer matter. Rovelli is right that all of that should be retired. What survives is the structural fact that any frame requires a frame-bearer the frame does not describe, and that this fact, far from being mystical, is the consequence of taking seriously what it means to be a finite physical system trying to give an account of the world from within it.

The right monism is therefore not Rovelli's flat physicalism — “the soul is of the same basic nature as a thunderstorm” — but a stronger one, in which both the soul and the thunderstorm are appearances within a single reality whose nature is not exhausted by either category. Contemplative traditions have a vocabulary for this: they call the appearance-side of reality the world of name and form, and the reality-side something that cannot be objectified because it is what makes objectification possible in the first place. The same structural insight is available without that vocabulary. It is just the observation that the standpoint that gives accounts cannot itself be fully accounted for from within.

This is not a return to dualism. It does not multiply substances. It refuses, instead, to grant priority to the third-person description at exactly the point where Rovelli's own perspectivism should refuse it. And — uniquely among non-dual readings of physics — it stages a falsifiable claim. If the bandwidth-dependent visibility experiment confirms what the framework predicts, the structural asymmetry between knower and known is doing physical work at a measurable scale. If it doesn't confirm it, the disagreement returns to philosophy.

## Same Direction, Different Distance

Rovelli ends his essay generously: “We have souls. We have an inner self. We can treat ourselves as transcendental subjects in the Kantian sense.” I would echo this without qualification. The disagreement is not about whether we have an inner life. The disagreement is about what that fact tells us about the structure of reality.

His argument moves correctly in three steps and then stops. Reject Cartesian dualism. Refuse the view from nowhere. Insist that knowledge is embodied. These three steps are also mine. But his premises imply a fourth step his conclusion declines to take. If any account is from within, and if every attempt to close the epistemic–ontic circle produces a further appearance, then the right response is not to grant the third-person picture quiet priority. It is to recognize that the picture and the picturer co-arise, and that no picture can finish itself by including its own author.

The hard problem in Chalmers's substance-dualist form is dissolved. Rovelli is right about that. What he misses is that something the hard problem was pointing at — the asymmetry between the describing standpoint and the described — survives the dissolution. Honouring that survival doesn't require a second substance. It requires a different kind of monism, with a hook into experiment.

That, I want to argue, is the move Rovelli's own argument should have made.

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

*"The eye that mistakes itself for the world cannot describe the eye."*