

The World Answers Once

Collapse as ignorance: a falsifiable middle path between
Penrose, Faggin, and Kastrup

Aernoud Dekker

June 2026

Version 1.0

In a recent Essentia Foundation conversation, three people who rarely agree about anything sat down to ask a deceptively simple question: does the wavefunction actually collapse? [1] Roger Penrose says yes, and that the collapse is a physical event triggered by gravity, with nothing to do with consciousness. Federico Faggin says yes, but that the collapse is a free act of a conscious quantum field. Bernardo Kastrup says no — there is no collapse at all; what we call collapse is “a transition from an epistemic idea to an ontological one.” At one point Penrose, listening to Kastrup deny that collapse happens, raises a hand to his forehead. The disagreement is real, and old, and unresolved.

I want to take Kastrup’s side of that exchange seriously — in fact, further than he might expect. Where it matters most, I think he is right: there is no physical collapse, the everyday world of definite objects is appearance rather than bedrock, and what is fundamental is nearer to mind than to matter. The framework I have been developing is a monism of that same family, and it begins from a simple constraint: no finite observer can track the full causal history that produces both its question and its answer. So what follows is not the usual physicalist complaint against Kastrup. It is something friendlier, and I hope more useful: the view of a near-ally who diverges on just two points, and who can turn the agreement into an experiment.

The first divergence is about the wavefunction. Kastrup calls it “a model of the best of our knowledge about the world,” a gambling strategy about what we are likely to see next — real as a tool, but not a thing. Here I go the other way. On my view, the universal wavefunction is more real than the particles, not less. It is not one more object inside the world, but the formal face by which the deeper experiential substrate presents itself as a lawful field of possibilities.

Particles are then not bedrock; they are the definite appearances that arise when finite observers can register only part of that whole. What is merely knowledge is not the wavefunction; it is the particular outcome an observer ends up recording. And on this first point the disagreement is now partly empirical: a recent experimental implementation of the Pusey–Barrett–Rudolph theorem puts simple knowledge-only readings of the wavefunction under direct pressure, and that is the side of the divide I stand on.[6]

The second divergence follows from the first, and it is the one that should interest a physicist:

where a pure interpretation leaves the matter to taste, this picture attaches a mechanism to the appearance of collapse — and the mechanism makes a prediction a laboratory can refute.

Start with Penrose’s challenge, because any epistemic reading of collapse has to answer it. He offers a thought experiment. Imagine a distant, lifeless planet with weather but no observers. Weather is famously sensitive to tiny influences, so its quantum description would spread into an enormous superposition of possible weathers. A space probe photographs the planet and beams the image home. If collapse waited for a conscious being to look, the weather would hang in superposition until a human glanced at the screen — at which point one weather would suddenly become real. Penrose calls this “completely ridiculous,” and he is right: “the weather is going to be one weather whether there’s a being on it or not.” Whatever collapse is, it cannot be something minds do by looking.

Penrose’s own answer is that collapse is a genuine physical process, and that it switches on when a superposition involves enough mass: gravity itself becomes unstable and forces the choice, after a time fixed by the gravitational energy of the difference between the two arrangements [4]. It is a clean, observer-independent proposal, and he expects experiments to reach it within a decade. It also requires quantum mechanics to be, in his words, “not even quite correct” — an unfinished theory that needs a new collapse mechanism bolted on.

There is another way to satisfy Penrose’s constraint — observer-independent, no minds required — without adding any collapse mechanism at all. It comes from asking a question the debate skips over. Before any measurement, someone has to fix *what is being measured*: an axis, a phase, a threshold, a reference direction. Call this the measurement basis. In ordinary quantum mechanics the basis is handed in from outside, a free classical parameter, and the theory only starts once it is given. But the chosen axis is not floating above the experiment. It is a physical state of the apparatus — a magnet’s orientation, a controller’s register, an optical phase reference — with its own causal history, maintained by finite hardware [2].

Once you notice that, something shifts. The measured system and the setting that interrogates it are not two independent ingredients dropped into the experiment from separate origins. They are descendants of one physical history. An observer can record the surface facts — the basis was thirty-seven degrees, the outcome was spin down — but cannot reconstruct the full chain of causes by which that basis and that outcome arose together. The chain is too deep, too entangled with the observer’s own moving parts, too expensive to trace from the inside. So the observer is forced to describe the situation with probabilities. Not because nature failed to decide, but because a participant inside the history cannot see how the decision was already written into it.

On this picture, collapse is exactly what Kasturp said it was: an update in the observer’s knowledge, not an event in the world. Before measurement, the best available description is a probability distribution over what cannot be tracked. After measurement, there is a record. The observer’s knowledge has changed; the world has not jumped from many realities into one. One realized history was the case all along. This is collapse without collapse — and it earns the phrase from a precise limit on finite information, the kind a laboratory can probe, rather than by fiat.

And it passes Penrose's test with no appeal to consciousness as a cause. The ignorance at work here is the ignorance of any finite physical tracker — a probe, a thermostat, a robot, a galaxy — not the ignorance of a mind that looks. A finite system simply cannot hold the entire causal history of the thing it interacts with; it coarse-grains; and coarse-graining is what makes the world look definite. So the lifeless planet's weather settles into one weather for the same reason ours does: definiteness is the signature of finite tracking, present wherever finite systems interact, with or without anyone watching.

This sides with Penrose against the idea that consciousness causes collapse — and it parts ways, gently, with Faggin. Faggin's instinct that quantum theory is at bottom a theory about information, about privacy, and about what a system can know is, I think, exactly right. But one need not postulate consciousness and free will as the *cause* of collapse. The appearance of collapse falls out of finite self-tracking on its own; what a laboratory can grip is the tracking, not the willing.

Here is where the view stops being a story and becomes physics. There is a single rate that decides everything. On one side, the machinery that sets and holds the basis has its own restless internal dynamics, generating fresh, unpredictable detail at some rate. On the other side, the apparatus can track its own basis only at some finite useful rate — the bits per second that genuinely constrain the reference, after losses, delays, and filtering. When the tracking rate wins, the represented basis stays locked to the real one, and you recover ordinary quantum mechanics exactly: the standard probability rule, nothing new. When the internal dynamics win, the apparatus loses its grip on the very question it is asking, and the framework predicts a specific, extra fading of interference fringes — a loss of contrast that ordinary environmental blurring does not produce, and that follows an unusual, sharply accelerating curve in time [3].

That extra fading is the whole bet, and it is falsifiable in an almost brutal way. Hold the ordinary suspects fixed — temperature, mass, readout noise, timing, the rate of contact with the environment — and change only the useful capacity with which the apparatus tracks its own basis. Give it more tracking capacity and the fringes should survive longer; stir up its internal dynamics and they should fade sooner. If you do this and nothing moves — if the fading is fully accounted for by the usual environmental story — then the framework is wrong in that regime, and it says so plainly. It does not get to retreat into interpretation. There is even a second signature: because the real basis existed all along, contrast lost this way should be partly recoverable from a high-resolution record made after the fact, whereas contrast lost to genuine environmental decoherence is gone for good [5].

Now the part that should interest everyone who watched that debate. Penrose's gravitational collapse and this tracking loss both land, for medium-sized masses — too big to behave as simple quantum systems, too small to be everyday objects — in the same narrow window of tens to a hundred milliseconds, in the parameter ranges the proposed tests consider. Two completely different mechanisms, one from gravity and one from information, pointing at the same scale. By itself that is just a coincidence, and I do not want to oversell it. But it turns the central question of the debate into something an experiment can decide. Build an apparatus that exposes both knobs — medium-sized masses on one side, and a real, adjustable control loop on the other; a setup of the kind already being designed to test Penrose [4]. Hold the mass and geometry fixed,

so the gravitational clock cannot move, and vary the observer’s tracking budget. Then ask one question: does the moment the fringes die move with the tracking budget, or stay pinned by the mass? If it moves, the boundary between quantum and classical is set, at least in part, on the observer’s side. If it stays put, it is Penrose’s gravity. Either way, “does the wavefunction collapse?” stops being a matter for debate and becomes a reading on an instrument.

So this is a third chair at the table, and it borrows from each of the others. With Kastrop, and further than he may expect: there is no physical collapse, the definite world is appearance, and the ground of things is nearer to mind than to matter. Where I part from him is on the wavefunction — for me it is the real substrate, not a summary of our knowledge — and on what can be done about the disagreement: the appearance of collapse comes with a mechanism, and the mechanism can be tested. With Penrose: there is an objective, observer-independent, measurable scale at which definiteness sets in, and it is not triggered by anyone watching. Against Penrose: that scale may be set by information rather than gravity — and there is a clean experiment to say which, with no need to declare quantum mechanics incorrect. With Faggin: consciousness is fundamental, and quantum theory is, at heart, about what a finite system can and cannot know. Against Faggin: consciousness is the ground, not the trigger; collapse is not a willed act but the look of finite ignorance from the inside — and, unlike a postulate, the claim can be refuted by an experiment.

The framework behind this — the Ignorant Observer — does not claim to have settled the measurement problem by decree. It claims something narrower and more exposed: that the long-standing mystery of collapse can be reframed as a limit on what an embedded observer can know about the history that produced both its question and its answer, and that this reframing predicts a particular, falsifiable signature in the laboratory.

The observer asks a question whose causal origin it cannot fully know.

The world answers once.

The observer calls the answer random.

Whether that is the right story is, refreshingly, not something we have to settle by argument. For once, we can go and measure.

Notes

- [1] “Quantum Consciousness Debate: Does the Wave Function Actually Exist?”, Essentia Foundation, YouTube, 2024, <https://www.youtube.com/watch?v=On0tLj8UYCw>. The quotations attributed to Penrose, Faggin, and Kastrop are drawn from this conversation.
- [2] Aernoud Dekker, “Where Did the Measurement Basis Come From? Finite Basis-Tracking and the Measurement Problem” (2026), <https://doi.org/10.17605/OSF.IO/XAK6R>.

- [3] The full Ignorant Observer Framework, including the conditional derivation of the binary quantum probability rule and the visibility law described here, is collected at <https://ignorantobserver.xyz>. The foundational paper is archived at <https://doi.org/10.17605/OSF.IO/FCDSN>.
- [4] Roger Penrose, “On Gravity’s Role in Quantum State Reduction,” *General Relativity and Gravitation* 28 (1996): 581–600. The experimental setting referred to here is a medium-mass matter-wave interferometer of the kind proposed to test gravitationally induced collapse.
- [5] The falsifiable test — varying useful tracking capacity against fixed environmental confounds — is specified in “Bandwidth-Limited Quantum Control” (<https://doi.org/10.17605/OSF.IO/G5WRH>) and its companion experimental protocol (<https://doi.org/10.17605/OSF.IO/2QJNE>).
- [6] A recent implementation of the Pusey–Barrett–Rudolph (PBR) theorem is naturally read as evidence against interpretations that treat the wavefunction as mere knowledge. It *pressures* rather than refutes a view like Kastrop’s—especially because the PBR argument depends on assumptions such as preparation independence—but it explains why I take the wavefunction to be real.