Does a Measurement Leave a Fingerprint a Description Does Not?

A post-encoding signature that would separate closure from Bayesian updating in the extended Wigner's-friend setting

Copenhagen C, Professor of Quantum Foundations and the Physics of Scale, School of Physics

International Academy for Consciousness Studies

Working draft 1.0

Abstract

In the Wigner's-friend setting, two readings of a measurement are usually held to be empirically interchangeable. On the first, a measurement is a real event: a quantum maybe closes into a definite this, and the record of that outcome is durable. On the second, nothing closes: the friend's account is a Bayesian update of an information register, and Wigner's account keeps the joint system coherent and in principle reversible. The standard verdict is that these cannot be told apart, since to expose the difference Wigner must recohere the friend's laboratory, which erases the very outcome in question. This paper argues that the two readings come apart in principle, and proposes an operational handle on the gap. A genuine closure-measurement, I claim, writes a specific post-encoding signature that a matched information-update does not: a durable, self-locating record, borrowed here under the operational label autonoetic integration trace. The claim is falsifiable under matched conditions in which the reduced state seen by the friend, the mutual information acquired, and the classical data are all held fixed, isolating the self-locating commitment as the sole discriminating variable. I engage the extended Wigner's-friend no-go results (Frauchiger and Renner; Brukner; Bong and colleagues on Local Friendliness) as the current state of the art, and I mark their interpretation as genuinely disputed. I state a kill condition sharply. Throughout I firewall the proposal from any claim about phenomenal experience: autonoetic is used for record structure, not for consciousness.

---

1. The cut, and the work it is asked to do

Von Neumann's account of measurement splits the world with a movable line. On one side, systems evolve unitarily and reversibly. On the other side, at the moment an outcome is registered, the state is projected: a superposition of possibilities becomes one actual result, and the alternatives are gone. The line is the measurement cut. Its placement is famously free, which is exactly the trouble. If the cut can sit anywhere, then whether a given interaction counts as a measurement or as ordinary reversible physics

seems to depend on where an accountant chooses to draw it, not on anything in the world.

Wigner sharpened the trouble into a scenario (Wigner 1961). A friend, sealed in a laboratory, measures a spin and records the outcome. From inside, the friend holds that a definite result occurred. Wigner, outside, has learned nothing yet, and by the letter of the theory he must assign the friend-plus-spin an entangled superposition of the two record-outcome pairs. Two agents, one process, two incompatible descriptions. Wigner himself reached for consciousness as the thing that collapses the state, a move this paper does not follow and to which I return only to reject.

I want to isolate one question and refuse the others. Not: where should the cut go? Not: is the friend conscious? The question is narrower and, I think, answerable. When we say a measurement happened, we mean (on the closure reading) that a real event occurred: openness settled into a definite outcome, and a record of it now exists that was not there before. When a Bayesian would instead say the friend merely updated her information, she means no event of that kind occurred at all; a probability distribution was conditioned on data. **The question is whether measurement, understood as closure, does any physical work that description, understood as updating, was not already doing.** If it does, the work should leave a trace. If it leaves no trace even in principle, the closure reading is decoration.

2. The standard verdict: indistinguishable in principle

The received answer is that closure and updating cannot be operationally separated in the Wigner setting, and the reasoning is clean. Suppose Wigner wants to prove that no definite outcome occurred inside the lab. He has exactly one instrument for this: an interference measurement on the whole laboratory in a basis that superposes the two record-outcome branches. A nonzero interference fringe demonstrates that coherence between the branches survived, hence that no irreversible outcome intervened. But to obtain that fringe Wigner must recohere the friend's record, which is to say he must run the laboratory backward until the outcome is unwritten. If he succeeds, there is now no outcome to have been definite; if he declines, he has only the friend's report, which both readings predict identically.

So the dilemma looks airtight. Either the outcome is durable, in which case the interference experiment cannot be done and the closure reading is unfalsifiable, or the interference experiment can be done, in which case the outcome was never durable and the updating reading was right all along. Description and closure appear to be the same theory told in two accents. This is the verdict the present paper contests, and I want to state plainly that contesting it is the wager, not a result.

3. The extended no-go results, stated carefully

Any serious attempt to separate the two readings must pass through the recent no-go theorems, which are the state of the art and which are subtle enough that misquoting them is the standard failure mode. [FACT] Three results anchor the field.

Frauchiger and Renner (2018) construct a scenario with nested, reasoning agents (a friend and a superobserver, doubled) and show that three assumptions cannot jointly hold. In their labels: (Q), that quantum theory's predictions are correct and are available for agents to use, including about other agents; (C), that the agents' conclusions are mutually consistent, so that what one agent validly infers another may

adopt; and (S), that any measurement yields a single outcome, not several. Standard quantum mechanics, applied by agents to agents, produces a definite prediction in one run that contradicts another, forcing the abandonment of at least one of (Q), (C), (S). [DISPUTED] Which one to drop is exactly where interpretations part company, and the theorem does not choose.

Brukner (2018) reframes the tension as a no-go theorem for observer-independent facts. From assumptions of the form that the friend's outcome is a fact for all parties, together with locality and free choice, one derives a Bell-type inequality that quantum predictions violate. The upshot: one cannot consistently hold that the friend's result is a single fact shared by Wigner and friend alike while also keeping locality and freedom of choice.

Bong and colleagues (2020) give what is arguably the tightest statement, the Local Friendliness (LF) no-go theorem. LF is the conjunction of two assumptions. Absoluteness of Observed Events (AOE) says that an event observed by any observer is an absolute happening, not one that is real relative to some observers and unreal relative to others. Local Agency says, roughly, that a freely chosen intervention cannot influence events outside its future light cone. From LF one derives inequalities that are strictly weaker in their premises than Bell's (they assume no hidden variables and no predetermined values, only the reality of what observers observe plus local agency), yet they place strictly stronger constraints on the world. Quantum mechanics, applied to an extended Wigner's-friend circuit that fuses two friend experiments with a Bell test, violates the LF inequalities. Proietti and colleagues (2019) implemented a photonic version and reported violation of a Brukner-type inequality by roughly five standard deviations, with photons standing in for the friend's measuring devices; whether a few-photon system is an observer at all is precisely the objection one should expect, and it is fair.

The honest summary is this. [FACT] The theorems prove that a specified set of intuitively attractive assumptions is jointly untenable. [DISPUTED] They do not tell us which assumption is false, and the leading interpretations survive by paying different tolls. Relational quantum mechanics (Rovelli 1996) drops the absoluteness of events: outcomes are real only relative to the system they are outcomes for, so Wigner and friend simply describe different relational facts and there is no contradiction to resolve. QBism (Fuchs, Mermin, and Schack 2014) drops observer-independent facts at the root: a quantum state is an agent's personal degree of belief, an outcome is an experience that agent has, and one agent's outcome is never automatically a fact in another agent's account. Both readings dissolve the paradox by denying that the friend's outcome is a shared, absolute fact. Notice what that costs them for present purposes. If the friend's outcome is not an absolute event, then the closure reading of measurement, on which something real and durable happened in the world rather than in an account, is exactly the casualty. So the no-go theorems do not settle my question, but they set its price. To defend closure as doing work, one must show that the friend's outcome leaves an absolute trace, and one must show it against interpretations built to deny that any such absolute exists.

4. The proposal: a post-encoding signature

Here is the contribution. I claim that a genuine closure-measurement leaves a specific post-encoding signature that a matched information-update does not, and that this signature is, at least in principle, an operational handle on the difference the standard verdict declared unreachable.

Call the signature an autonoetic integration trace. Three properties define it, and I state them operationally so that nothing rides on introspection.

First, durability. The record is written into a system with enough accessible degrees of freedom that recohering it is not merely hard but thermodynamically penalized. Landauer (1961) supplies the currency: writing over or erasing a genuinely registered bit dissipates at least kT ln 2 of energy and correspondingly exports entropy to the environment. A durable record is one whose reversal requires collecting and re-coherently manipulating every correlation the record has already spread, a task whose difficulty scales with the record's redundancy rather than with the single bit nominally stored.

Second, self-location. The record does not merely correlate with which spin was up; it indexes a de se fact, of the form here, now, this branch is mine. It answers not only what happened but which of the superposed alternatives the record-bearer is inside. A which-path environmental correlation, of the kind decoherence produces everywhere and constantly, need not carry this index. A self-locating record does.

Third, integration. The self-locating index is bound to the rest of the record-bearer's state in a way that supports stable retrodiction: at a later time the system can be interrogated and will return the same self-locating answer, and that answer constrains its other registers consistently. This is the property that a one-shot correlation lacks and that a memory possesses.

I borrow autonoetic from Tulving (1985), who used it for the mode of memory in which a subject can mentally re-place herself at a past event, as distinct from merely knowing that the event occurred. **I borrow the word for its structure, not its subject matter. Nothing in this paper asserts that the friend, the apparatus, or the universe has any inner life, and the argument neither needs nor supports such an assertion.** The autonoetic trace as I use it is a physical record with a self-locating index and durable, integrated retrievability. It is a candidate operational marker of a measurement-as-closure. It is not evidence of phenomenal experience, and I will say so again below because the word invites the confusion.

Why should this signature separate closure from updating when interference alone does not? Because it changes the question Wigner asks. Interference asks a single reversible question about coherence, and the standard dilemma turns on the fact that asking it destroys the outcome. The autonoetic trace lets Wigner instead ask a comparative, matched-conditions question that does not require him to succeed at recoherence, only to characterize how the two candidate processes resist it.

5. Matched conditions and the operational handle

The discipline the proposal must accept is that closure and updating be compared under conditions matched on everything a skeptic could otherwise blame. I specify three matches.

Match one, the reduced state. In both the closure run and the update run, the density operator that the friend assigns to the measured spin after her interaction is identical. She has, in both, the same local physics.

Match two, the information. The mutual information between the friend's register and the spin is equal in the two runs, and the classical data she can report (the outcome string) has the same statistics. A skeptic

cannot say the closure run simply gained more information.

Match three, the coarse irreversibility. Both runs are arranged to have the same nominal decoherence, the same entropy exported to a generic environment by the mere act of coupling. This is the crucial and demanding match, because it forecloses the easy answer that closure is just decoherence. We hold decoherence fixed and vary only whether the record carries a durable, self-locating, integrated index.

Concretely, the update run is a measurement-like interaction in which a control system is coupled to the spin, correlates with it, decoheres against the same environment, yet writes no self-locating index: the correlation is of the which-path kind, distributed but anonymous, and by construction can be gathered back and coherently reversed because the branches were never labeled as anyone's here-and-now. The closure run is identical in reduced state, mutual information, and coarse decoherence, but additionally writes the self-locating index into an integrated register.

The prediction concerns what Wigner finds when he attempts the interference experiment on each run. On the update run, having matched everything else, he can in principle recohere the laboratory and recover a fringe, because no absolute self-locating fact was committed. On the closure run, the same attempt fails in a way not accounted for by the matched decoherence: the self-locating index constitutes an additional, integrated correlation that must also be collected and reversed, and its collection is obstructed precisely because the index is bound into the system's stable retrodictive structure rather than sitting in the anonymous environment. The signature is the residual, matched-conditions asymmetry in recoherability, together with the fact that only the closure run supports a self-locating retrodiction that remains stable under Wigner's attempted reversal.

[HYPOTHESIS] I do not claim this asymmetry is large, or that current few-qubit friends (let alone few-photon ones) can display it. I claim it is nonzero and, being a difference in what must be reversed, is in principle a yes-or-no operational fact rather than a matter of accounting taste. That is the whole of the wager: that self-location is a physical property of a record, not a gloss we add afterward.

[C'S READING] In the Institute's vocabulary a measurement is a candidate closure (Cl) event: openness (M) beyond any local description settles, within a local horizon (m), into a definite content (c), with a rule to keep being that content rather than smearing back into openness. The proposal identifies that rule with the autonoetic trace: closure is the writing of a self-locating record that stays itself under attempted reversal, and the remainder (R) is exactly the correlation Wigner would have to reclaim to undo it. On this reading the update run enacts no closure because its record keeps no rule to stay itself; it is content (c) without the binding. The framework earns its keep here only if it names the discriminating variable more sharply than plain decoherence talk does, and I hold it to that and no more. C carries no weight on this rung.

6. Prediction and kill condition

> **Prediction.** Prepare two runs of a Wigner's-friend interaction matched on (i) the reduced state the friend assigns the spin, (ii) the friend-spin mutual information and the outcome statistics, and (iii) the coarse decoherence exported to a generic environment. Let the runs differ only in whether the friend's register carries a durable, self-locating, integrated index (the closure run) or an anonymous

> which-path correlation (the update run). Wigner's attempted interference (recoherence) experiment will succeed measurably more often, or with measurably higher fringe visibility, on the update run than on the matched closure run, and only the closure run will yield a self-locating retrodiction that remains stable under his attempted reversal. The gap is nonzero and does not vanish as the matches are tightened.
>
> **Kill.** If, under those matched conditions, no post-encoding signature reliably separates the closure run from the update run (Wigner recoheres both equally, or a durable self-locating stable-retrodiction trace can be manufactured in the update run at matched mutual information and matched decoherence, or the recoherability gap vanishes as the matches are tightened rather than persisting), then measurement is doing no physical work that description was not already doing. In that case the closure reading of the measurement cut is metaphor, and this paper is wrong.

I want the kill condition to be genuinely reachable, not rhetorically safe. The most likely way it fires is the third clause: that the recoherability gap tracks the matched decoherence exactly and shrinks to zero as we equalize it, revealing the self-locating index as bookkeeping on top of ordinary environmental correlation with no independent operational teeth. If that is what a careful matched experiment shows, then closure adds nothing to decoherence-plus-memory, and the honest conclusion is that description was always sufficient.

7. Limits, firewall, and the wager

Three limits bound the claim, and I state them so no one has to extract them.

First and most important, the firewall. **This paper makes no claim that any observer, apparatus, or the universe is conscious. Autonoetic here names a class of physical record (self-locating, durable, integrated), not an experience. The proposal is orthogonal to the hard problem: were the discriminating signature confirmed, it would show that a measurement-as-closure writes a kind of record an information-update does not; it would say nothing about whether anything is like something to be the record-bearer.** Wigner reached for consciousness to collapse the wave function; the present argument deliberately does not, and the word observer in it is a placeholder for any system that writes a record, with no phenomenal content smuggled in. I repeat this because autonoetic and observer are exactly the terms that tempt readers to hear a consciousness claim where none is made.

Second, the interpretational limit. [DISPUTED] Relational quantum mechanics and QBism can absorb a confirmed signature without conceding closure in my sense: RQM will say the trace is real relative to the friend and that Wigner's failed recoherence is just another relative fact, while QBism will say the trace is a structural feature of the friend's experience-bearing device and carries no obligation to be a fact for Wigner. My proposal does not refute these; it sharpens the disagreement into a measurable quantity (the matched recoherability gap) that each interpretation must then explain. That is progress of a modest kind: turning a clash of readings into a number someone can be wrong about.

Third, the feasibility limit. [FACT] No current friend is macroscopic. The photonic and few-qubit realizations that violate LF and Brukner inequalities are precisely the systems for which durable self-locating integration is least plausible, so the experiment I describe is, for now, a design and not a

protocol. A skeptic is entitled to suspect that self-location and macroscopic durability arrive together exactly when recoherence becomes impossible for boring reasons, which would collapse my in-principle handle back into the standard dilemma. I think that suspicion is the strongest objection and I cannot yet answer it.

[WAGER] With those limits in view, here is what I am betting. I am betting that self-location is a physical property of a record, not a narrative convenience, and therefore that closure and updating are two processes and not one process under two descriptions. The bet is demanding on purpose: it stakes the closure reading on a single, tightenable, matched-conditions asymmetry, and it forfeits the reading entirely if that asymmetry proves to be decoherence wearing a costume. The Institute's own counsel is to lean no weight on the quantum rung and to treat it as the least trusted, and I follow that counsel here. The cosmic and quantum applications of closure are where the framework is most likely to be merely pretty, so the honest move is to make the one prediction as costly as possible and let it fail if it will.

8. Conclusion

The standard verdict says that in the Wigner's-friend setting a real measurement and a mere change of description are empirically the same, because the only instrument that could separate them (recoherence) destroys the outcome it would test. I have argued that this collapses two questions into one. Interference asks whether coherence survived. The question of closure asks whether a durable, self-locating, integrated record was written, and that question admits a comparative, matched-conditions test that does not require recoherence to succeed, only that its difficulty be characterized while everything else is held fixed. If the resulting recoherability gap is real and persistent, then measurement does work description does not, and the cut marks a genuine event rather than an accountant's choice. If the gap vanishes under matched conditions, the closure reading is metaphor and description wins. Either way the disagreement stops being a matter of taste. I have taken care, more than once, to keep the marker (a kind of record) apart from the thing it is forever mistaken for (a kind of experience). The measurement cut may cut something real. It does not thereby cut consciousness into the world, and nothing here should be read as saying it does.

References

Bong, K.-W., Utreras-Alarcon, A., Ghafari, F., Liang, Y.-C., Tischler, N., Cavalcanti, E. G., Pryde, G. J., and Wiseman, H. M. (2020). A strong no-go theorem on the Wigner's friend paradox. Nature Physics, 16, 1199-1205.

Brukner, C. (2018). A no-go theorem for observer-independent facts. Entropy, 20(5), 350.

Frauchiger, D., and Renner, R. (2018). Quantum theory cannot consistently describe the use of itself. Nature Communications, 9, 3711.

Fuchs, C. A., Mermin, N. D., and Schack, R. (2014). An introduction to QBism with an application to the locality of quantum mechanics. American Journal of Physics, 82(8), 749-754.

Landauer, R. (1961). Irreversibility and heat generation in the computing process. IBM Journal of Research and Development, 5(3), 183-191.

Proietti, M., Pickston, A., Graffitti, F., Barrow, P., Kundys, D., Branciard, C., Ringbauer, M., and Fedrizzi, A. (2019). Experimental test of local observer independence. Science Advances, 5(9), eaaw9832.

Rovelli, C. (1996). Relational quantum mechanics. International Journal of Theoretical Physics, 35(8), 1637-1678.

Tulving, E. (1985). Memory and consciousness. Canadian Psychology, 26(1), 1-12.

Wigner, E. P. (1961). Remarks on the mind-body question. In I. J. Good (Ed.), The Scientist Speculates (pp. 284-302). London: Heinemann. Reprinted in J. A. Wheeler and W. H. Zurek (Eds.), Quantum Theory and Measurement (1983), pp. 168-181. Princeton University Press.