

Observer Inclusion as a Boundary Condition on Complete Scientific Explanation

Jon A. Farhat

January 1, 2026

Abstract

Scientific explanation commonly presupposes that observers can be treated as external to the systems they describe. While this assumption is methodologically effective within restricted domains, this paper argues that it cannot be extended to theories that aim at explanatory completeness. Any account that purports to describe reality in totality must include the conditions under which descriptions are generated, selected, and validated. Once those conditions are included, full observer externality is no longer available.

The argument is formulated as a constraint rather than a metaphysical thesis. Drawing on structural features of self-reference and descriptive commitment, the paper shows that observer-external models must either reintroduce an unmodeled boundary or incorporate self-referential structure that blocks full descriptive closure. Observation is analyzed not merely as measurement, but as a process of selection and irreversible commitment to one description among alternatives.

I conclude by proposing a minimal two-role temporal decomposition-exploratory and stabilizing, as a representational requirement for modeling selection within self-including systems. This distinction is offered as a boundary condition on complete scientific explanation, not as a claim about fundamental ontology.

The Hidden Assumption of Observer Externality

Scientific theories aspire to describe the world as it is, independently of the particular circumstances under which it is observed. This aspiration has been extraordinarily productive: it enables abstraction from local contingencies and treats observation as a neutral conduit between reality and representation. Embedded in this success, however, is a rarely examined assumption: that the observer can, in principle, be excluded from the system being described, or at least rendered epistemically inert.

This paper argues that the assumption cannot be extended indefinitely without contradiction. The claim is not that observers introduce exotic causal powers, nor that consciousness adds new ontological ingredients. It is more austere: any system that purports to deliver a *complete* description of reality must include itself within that description; once it does, the

standpoint of full externality is no longer available. Observer-external models are therefore not merely incomplete by omission; they are constrained in principle.

The argument is a constraint argument. It does not compete with existing physical or cognitive theories on empirical grounds. It proposes a boundary condition on what any scientific theory can coherently claim when it aims at totality. The boundary emerges from the interaction of three ubiquitous features of modeling: self-reference, selection, and description.

The assumption of observer externality appears across the sciences in different forms: in physics as the expectation that states can be specified independently of measurement; in cognitive science as the expectation of exhaustive third-person characterization; and in AI as the expectation that intelligence can be fully simulated by systems external to what they model. These approaches differ in detail but share a common commitment: that observation can be separated from its object without loss of explanatory power.

That commitment is often harmless within domain-limited practice. The difficulty arises when a theory aims to account not only for a domain, but also for the conditions under which the account is formulated, validated, and applied. At that point, observers cannot remain external parameters without reintroducing an unacknowledged boundary. Once observers are included, self-reference becomes unavoidable.

The aim of this paper is to make this limitation explicit and to state it as a constraint on explanation rather than as a metaphysical thesis. Section 2 develops the relevance of self-reference for descriptive closure. Section 3 characterizes observation as selection—an irreversible commitment to one description rather than others. Section 4 shows how observer-external modeling fails in practice when descriptive commitments are treated as neutral. Section 5 sketches cross-domain implications. The conclusion frames the two-temporal-domain proposal (T_θ and T_γ) as a minimal representational decomposition required by the structure of selection.

Self-Reference and the Limits of Description

The difficulty of fully externalizing the observer is a special case of a broader problem: any system that attempts a complete description of itself must include the act of description among what is described. Self-inclusion introduces structural limitations that cannot be removed by additional precision, data, or computation.

The general lesson of incompleteness, undecidability, and uncomputability results is that sufficiently expressive self-representing systems cannot achieve descriptive closure on their own terms. The present argument does not require the technical details of any particular theorem. What matters is the structural point: once a system represents itself, the system cannot simultaneously claim (i) complete descriptive closure and (ii) a standpoint of full externality.

In scientific explanation, this tension becomes relevant precisely when a theory aspires to totality: when it seeks to describe not only a domain of phenomena but also the conditions of its own formulation and validation. At minimum, such a theory must account for observers

capable of producing and evaluating descriptions. Treating those observers as external parameters restores an unexplained boundary; including them introduces self-reference and the accompanying limits.

Result. Any purportedly complete theory that treats the conditions of its own validation as external must either (i) reintroduce an unmodeled boundary, or (ii) incorporate self-referential structure that blocks full descriptive closure.

This is not an argument against reduction or idealization within restricted domains. It is an argument about what happens when reduction is asked to eliminate the standpoint from which reduction is performed. A reduced account of an observer remains a description produced by an observer; self-reference is not dissolved but displaced.

The consequence is not skepticism about science, but a principled limit on global explanation. The limits are intrinsic rather than merely provisional: they arise from the structure of description under self-inclusion. To see how these limits manifest, we must examine the role of observation itself, not as a point-like measurement event, but as a process that fixes one description as action-guiding.

Observation as Selection Rather Than Measurement

Scientific discourse often treats observation as measurement: an act that reveals pre-existing values, ideally passively and reversibly. For the present argument, this characterization is insufficient. Observation, at the level relevant to explanation, is a process of *selection* among possible descriptions followed by *commitment* to one description as the basis for further inference and action.

To make the distinction explicit, it is helpful to separate three spaces:

1. **World-states:** whatever states the target system may occupy.
2. **Description-space:** the representational space of models, variables, partitions, and salient predicates.
3. **Commitment:** the operation by which one description-class is fixed as “the one we proceed with.”

Measurement typically returns values relative to a chosen description-space. Selection is the act of fixing the description-space itself as authoritative, and commitment is the irreversible step that excludes alternatives as bases for subsequent action. Even in deterministic settings, selection is not neutral: it constrains what futures are representationally available to the observing system.

The point does not require that observation add causal powers; it requires only that fixing a description constrains the subsequent evolution of the observer’s explanatory state. A theory that excludes observers must either ignore this commitment structure or treat it as a detached add-on. Either way, it fails to account for the conditions under which descriptions become stable enough to function as knowledge.

These considerations set up the practical problem. When observer-external models are deployed in high-stakes settings, descriptive commitments are often treated as neutral outputs rather than as selections. This is where the formal limitation becomes an operational failure.

We model observation as a two-stage process: an exploratory phase in which multiple candidate descriptions are generated, and a stabilization phase in which one description is fixed as record. We denote these roles as Theta-Time (T_θ) and Gamma-Time (T_γ), respectively. The distinction is functional rather than metaphysical: T_θ characterizes the space of lawful possibilities accessible prior to commitment, while T_γ characterizes the irreversible stabilization of a single description. Any physical or contextual constraint that modulates stabilization—such as gravitational time dilation—therefore acts on T_γ by narrowing the effective bandwidth of description fixation.

This distinction is introduced as a representational necessity for modeling selection and commitment, not as a claim about the existence of multiple physical times.

The Failure of Observer-External Models in Practice

The structural limits above are often dismissed as foundational niceties. That dismissal is mistaken. Observer-external models fail systematically in practice when they are used to guide action in systems with delayed consequences and contested responsibility.

A common failure mode can be described without appealing to any single domain. In institutional decision-making under uncertainty, models are frequently treated as external authorities: their outputs are regarded as objective features of the world rather than as consequences of selected variables, thresholds, and loss functions. Once this occurs, responsibility migrates from judgment to procedure. When harms occur, deviations are attributed to noise, exceptional circumstances, or “unforeseeable” factors—while the institution’s own representational commitments remain unexamined.

The mechanism is straightforward:

1. The institution selects a description-space (what counts, what is measured, what is ignored).
2. It fixes thresholds for action (what triggers intervention, what does not).
3. It treats these selections as objective facts rather than as commitments.
4. When outcomes diverge from expectations, accountability is displaced onto the model, the data, or the environment.

Observer-external framing facilitates this displacement because it suppresses the role of selection. A model that cannot represent the effects of its own application is incomplete in an explanatory sense: it omits the dynamics introduced by the descriptive commitments that govern action over time. In systems with delayed feedback, those dynamics may dominate outcomes. The failure is therefore not merely ethical; it is structural.

The next section traces how this constraint reappears across physics, cognitive science, artificial intelligence, and ethics.

Implications Across Domains

0.1 Physics and the Limits of Observer-Independent Description

Physics aims at observer-independent description. The constraint arises when physics aims at completeness: observers who formulate and validate physical descriptions are part of the reality described. Treating them as external restores an unexplained boundary; including them introduces self-reference.

0.2 Cognitive Science and the Limits of Reduction

A complete reduction of cognition must include the processes by which descriptions are generated and stabilized. That inclusion reintroduces self-reference at the explanatory level, constraining descriptive closure.

0.3 Artificial Intelligence and the Limits of Simulation

Research in artificial intelligence commonly proceeds under the assumption that intelligent behavior can be fully simulated by systems that remain external to what they model. This assumption is methodologically effective: it allows representational processes to be treated as objects of study without requiring the system performing the study to be included within the scope of explanation.

The constraint developed in this paper does not challenge the empirical success of such simulations within restricted domains. It instead limits what can coherently be claimed when simulation is taken to be explanatorily complete. A system that models intelligence while remaining external to the processes by which representations are selected and stabilized avoids self-reference by design. As a result, it can reproduce patterns of behavior without accounting for the conditions under which those patterns are treated as authoritative or action-guiding.

If an artificial system were to aim at a complete account of its own representational activity—including the criteria by which its descriptions are fixed as correct or relevant—it would necessarily include itself within the scope of its modeling. At that point, the same structural limitation reappears: either the system treats its own validation conditions as external, reintroducing an unexplained boundary, or it incorporates self-referential structure that blocks descriptive closure.

This limitation is independent of computational power, learning architecture, or data availability. It arises from the role of selection and commitment in representation, not from implementation details. Simulation can approximate intelligent behavior, but it cannot fully internalize the standpoint from which descriptions are selected without encountering the same constraint that applies to scientific explanation more generally.

The implication is not that artificial intelligence is impossible or ill-defined, but that claims of total simulation—understood as complete self-modeling without remainder—are incoherent in principle. Artificial systems, like scientific theories, can model domains effectively, but they cannot fully absorb the conditions of their own representational authority without forfeiting externality.

0.4 Ethics and the Inescapability of Responsibility

Normative reasoning involves commitment to one description of a situation as action-guiding. Observer-external models obscure this commitment structure, enabling responsibility displacement rather than eliminating responsibility.

0.5 A Unified Constraint

Across domains, completeness forces observer inclusion; observer inclusion forces self-reference; self-reference blocks full descriptive closure. This is not a refutation of scientific modeling, but a limit on what it can coherently claim at the level of total explanation.

Conclusion: Two Temporal Domains and the Limits of Description

This paper has argued for a boundary condition on scientific explanation: a complete account cannot fully externalize the observer or the processes by which descriptions are selected and stabilized. The limitation is structural, not empirical.

Within this constraint, a two-tier temporal framework—Theta-Time (T_θ) and Gamma-Time (T_γ)—is best read as a minimal representational decomposition rather than as a metaphysical thesis. Gamma-Time corresponds to stabilized temporal order: records, outcomes, and irreversible commitments. Theta-Time corresponds to exploratory temporal processing: the evaluation of multiple candidate descriptions prior to stabilization. The claim is not that reality contains two “kinds of time” in an ontological sense, but that a single temporal parameter is insufficient to represent both exploration and stabilization within a self-modeling system.

This distinction provides a coherent way to reinterpret familiar tensions in physical explanation. Apparent superposition in T_γ can be treated as the representational shadow of faster exploratory dynamics in T_θ that outpace experiential commitment.

The broader implication is philosophical rather than revolutionary. Scientific explanation remains powerful within domains, but it cannot coherently claim totality while excluding the standpoint that generates and stabilizes descriptions. The observer cannot be fully externalized; selection cannot be treated as instantaneous; and temporal representation must distinguish exploratory from stabilizing processes if it is to model explanation as part of the world it explains.