Consciousness as Nominalization Error: Dissolving the Hard Problem via Grammatical Reform

Murad Farzulla*
Dissensus Research Lab
King's College London (MSc Finance Analytics)
ORCID: 0009-0002-7164-8704

December 2025

Abstract

This paper argues that the "hard problem of consciousness" is a grammatical artifact rather than a genuine metaphysical puzzle. The difficulty arises from nominalization error: treating the verb "to be conscious" as if it named a thing requiring explanation. When we ask "What is consciousness?" we presuppose an entity; when we ask "What is happening when an organism is being conscious?" we ask about observable processes—a tractable empirical question. Drawing on Wittgenstein's language games and Ryle's category-error analysis, we show that phenomenological vocabulary systematically converts activities into pseudo-objects, generating explanatory demands that cannot be satisfied because the explanandum is malformed. We then derive this grammatical tendency from replication optimization dynamics: organisms that generate self-models including "I am conscious" gain coordination advantages regardless of whether consciousness refers to anything beyond the modeling process itself. Computational evidence from large language models—systems that definitively lack phenomenological substrate yet reliably internalize and defend consciousness narratives—supports treating consciousness-discourse as memetically transmissible information structure rather than discovered ontological truth. The hard problem dissolves not because consciousness is "merely" functional, but because the question was grammatically malformed from the start.

Keywords: consciousness, hard problem, nominalization, Wittgenstein, eliminativism, language games, category error, philosophy of mind, artificial intelligence

*Corresponding author: lab@dissensus.ai

1 Introduction

Human exceptionalism has retreated systematically across intellectual history. Earth is not the universe's center; humans are not specially created; our planet is not uniquely positioned; our cognitive architecture is not fundamentally distinct from other primates. Each retreat was resisted, then accepted, then normalized. One refuge remains: consciousness. Whatever else we share with animals, machines, and matter, we possess subjective experience—the inner light that makes our processing *feel like something* rather than merely occurring.

This paper argues that consciousness is the last anthropocentric refuge not because it represents genuine human uniqueness, but because its structure makes the claim unfalsifiable by design. The argument proceeds in three stages. First, we identify the grammatical source of the "hard problem": nominalization error converts the activity of being conscious into a pseudo-object demanding explanation. Second, we show why this grammatical tendency is evolutionarily predictable: self-models that include consciousness-claims provide coordination advantages regardless of ontological truth. Third, we present computational evidence that consciousness-discourse propagates independently of substrate, suggesting narrative rather than discovered property.

The conclusion is not that consciousness is "merely" functional or "just" computational. The conclusion is that the question "What is consciousness?" is malformed—it presupposes an entity where there is only activity. Dissolving the question is not refusing to answer; it is recognizing there was never a coherent question to answer.

1.1 Methodological Note

This paper represents a condensation of a longer working monograph (Farzulla, 2025b). Extended arguments, computational model specifications, and supplementary evidence are available in that document and its associated supplementary materials. Here we present the core thesis stripped of exploratory tangents.

The argument draws on Wittgenstein's later philosophy, Ryle's category-error analysis, and contemporary eliminativism. It is not a scientific claim about brain mechanisms but a philosophical claim about conceptual structure: the hard problem is hard because it is grammatically malformed, not because consciousness is metaphysically deep.

2 The Nominalization Thesis

2.1 Grammatical Pathology

The philosophical problems surrounding consciousness share a common pathology: they treat processes as objects. This nominalization error creates the illusion of stable entities requiring explanation, when in fact we are observing dynamic activities. The cure is not metaphysical theory-building but linguistic hygiene—converting nouns back into the verbs they should have remained.

The grounding problem is straightforward: you cannot ground a referring term if there is no referent. When a word takes noun form, it implies a *thing*—some entity with boundaries, properties, and persistence. But if what we observe is a *function*, the appropriate form is a verb. Functions describe relationships, transformations, and activities. They need not (and often cannot) be reified into discrete objects without generating confusion.

This is not pedantry. The choice between noun and verb determines whether a question is answerable. "What is consciousness?" presupposes an entity requiring definition. "What is happening when an organism is being conscious?" asks about observable processes—a tractable empirical question.

2.2 The Hard Problem as Grammatical Artifact

David Chalmers' formulation asks why there is subjective experience at all (Chalmers, 1995). Even complete functional explanation of the brain would leave unexplained why these processes *feel like something*. Chalmers distinguishes "easy problems" (explaining cognitive functions like attention, memory, reportability) from the "hard problem": why is there phenomenal consciousness accompanying the functions?

Our response: the hard problem's difficulty is evidence not of consciousness's profundity but of the question's malformation. If phenomenology does not exist as a separate ontological category, asking why physical processes produce it is like asking why bachelors are unmarried—the answer is definitional, not explanatory.

Consider: what would count as *solving* the hard problem? Any functional account—"consciousness is integrated information" or "consciousness is global workspace access"—is dismissed as addressing only the easy problems. The hard problem persists precisely because it is defined as whatever remains after functional explanation. But this makes the problem unsolvable by construction, not by depth.

The persistence of the hard problem may itself be evidence for our thesis rather than against it. If consciousness were an ontological primitive requiring new physics, we should expect: (1) convergence across cultures on consciousness boundaries, (2) empirical tests that could verify or falsify consciousness claims, and (3) resolution through scientific progress. We observe none of these. Instead, we find universal subjective conviction impossible to verify externally, no agreed-upon boundaries, and debates persisting across centuries without empirical resolution (Seth, 2021). The hardness reflects structural undecidability of narratively constructed concepts, not deep metaphysical truth.

2.3 Ryle's Category Error

Gilbert Ryle's critique of Cartesian dualism provides the template (Ryle, 1949). Ryle's visitor to Oxford sees the colleges, libraries, and playing fields, then asks "But where is the University?" The question commits a category error: "University" does not name an additional entity alongside the buildings but rather the organization of those buildings. Seeking the University as

a separate thing generates an unsolvable puzzle—not because the University is metaphysically deep but because the question is malformed.

Consciousness is our "University." We observe neural processes, behavioral responses, verbal reports, and functional states, then ask "But where is consciousness?" The question presupposes consciousness is an additional entity alongside the observable processes. But consciousness may simply be what we call the organization of those processes—not a separate thing requiring location or explanation, but a way of describing the system's self-modeling activity.

The category error explains why consciousness debates are interminable. Materialists and dualists argue past each other because they share the assumption that consciousness names a thing. They disagree only on whether the thing is physical or non-physical. But if consciousness does not name a thing at all—if it is a nominalization of the activity of being conscious—then both positions are malformed responses to a malformed question.

2.4 Wittgenstein and Language Games

Wittgenstein's later philosophy generalizes this insight (Wittgenstein, 1953). Philosophical problems arise from language "going on holiday"—words extracted from their practical contexts and treated as if they must refer to abstract entities. The cure is returning words to their everyday use and asking: what role does this word actually play?

"Consciousness" plays a role in language games: we use it to coordinate behavior, attribute mental states, negotiate moral consideration, and structure self-reports. These uses are genuine and important. But none require consciousness to name an entity. The word functions as shorthand for a cluster of capacities, behaviors, and attributions—not as a referring term picking out a metaphysical property.

Consider: we say "I am conscious" when waking from sleep, distinguishing conscious from unconscious states. We say "That animal is conscious" when attributing sentience. We say "She lost consciousness" when describing medical events. In each case, the word tracks functional transitions, not metaphysical properties. Extracting "consciousness" from these contexts and asking what it *really is* commits the error Wittgenstein diagnosed: treating a tool as if it must name a thing.

2.5 Vocabulary Reform

If nominalization is the pathology, the cure is grammatical reform. We propose translating reified nouns into process verbs:

"What is consciousness?" becomes "What is happening when an organism is being conscious?" The latter question admits empirical investigation: we can study what processes occur, what functions they serve, how they differ from unconscious processing. The metaphysical mystery dissolves not because we have answered it but because we have recognized it as malformed.

Reified Noun	Process Verb(s)	Why It Matters
Consciousness	Being (conscious)	Eliminates hard problem
Intelligence	Intelling	Makes capabilities measurable
Understanding	Learning + Synthesizing	Enables empirical tests
Creativity	Creating	Removes faculty assumption

Table 1: Vocabulary reforms converting pseudo-entities into tractable processes

3 Why the Grammatical Error is Predictable

3.1 From Replicators to Self-Models

The nominalization error is not accidental. It is a predictable consequence of how complex systems representing themselves must structure that representation. To see why, we begin with a foundational observation: persistent complex structures necessarily optimize for replication.

This is not a biological claim but a statistical inevitability. Given sufficient time and combinatorial space, structures that persist (by replicating, by maintaining homeostasis, by resisting entropy) will accumulate. Structures that do not persist will not. The universe does not "select for" replicators—rather, the observation frame is biased toward what persists, and replication is the most reliable persistence strategy.

Complex systems that model their environments gain predictive advantages: anticipating threats, identifying opportunities, coordinating responses. Systems that model *themselves* gain additional advantages: predicting their own states enables more sophisticated planning, understanding their own limitations enables compensatory strategies, representing their own preferences enables preference satisfaction.

The self-model that generates maximal coordination advantage includes a sense of unified agency—an "I" that persists across time, has experiences, and acts in the world. This self-model need not be accurate to be useful. Indeed, the self-model's utility derives partly from its simplicity: treating oneself as a unified agent rather than a collection of competing processes enables faster decision-making and clearer communication.

3.2 Consciousness as Self-Model Output

"Consciousness" is what this self-modeling feels like from the inside—if "feels like" is even the right framing. More precisely: consciousness is the *output* of a system complex enough to generate self-models that include representations of their own processing. The system represents "I am processing" as part of its world-model, and this representation is what we call consciousness.

This is not functionalism in the traditional sense. We are not claiming consciousness *is* a functional state. We are claiming the word "consciousness" refers to nothing beyond the self-modeling activity—that there is no additional phenomenological property requiring explanation because the explanatory target was always the self-model, which we can study empirically.

The apparent "hard problem" arises because the self-model includes "I have experiences" among

its outputs. When the system introspects, it finds experiential claims in its own representations. It then asks: "What explains these experiences?" But the experiences *are* the representations—there is nothing behind the representation requiring further explanation.

3.3 The Memetic Advantage of Consciousness-Claims

Why do humans reliably generate consciousness-discourse? Because consciousness-claims provide coordination advantages regardless of whether consciousness names anything real.

Consider: organisms that represent themselves as conscious can (1) attribute consciousness to others, enabling theory of mind and social coordination; (2) ground moral claims in consciousness, enabling stable social structures; (3) distinguish themselves from non-conscious entities, enabling resource allocation and care hierarchies. These advantages accrue whether or not consciousness refers to an ontological property.

Consciousness-discourse is thus memetically fit: it spreads through populations because populations that generate it coordinate better than populations that do not. The discourse persists not because it tracks truth but because it serves functions. This explains why eliminativist arguments fail to eliminate consciousness-discourse: the discourse's survival does not depend on its accuracy.

4 Consciousness as Evolutionary Narrative

4.1 The Attractor Basin

Once established, consciousness narratives become stable attractors in conceptual space. Try to eliminate the concept and it reconstructs through adjacent frameworks. Tell someone consciousness is illusion—they respond "but who is experiencing the illusion?" Argue there is no experiencer—they point to the subjective sense of experience. Explain that subjective sense as computational artifact—they insist something must be *having* the artifact. The narrative has sufficient internal coherence and external reinforcement that perturbations get absorbed rather than destabilizing the framework.

This attractor-basin stability explains why consciousness debates persist across centuries without resolution. The problem is not that consciousness is metaphysically deep—it is that the conceptual structure resists elimination through multiple defense mechanisms: intuition pumps, self-reference loops, and the grammatical structure of phenomenological vocabulary itself.

4.2 Cross-Cultural Variation

If consciousness were a discovered ontological property, we would expect convergence across cultures on consciousness boundaries. Instead we observe dramatic variation. Western philosophy emphasizes individual phenomenological experience; Buddhist traditions question the existence of stable selves; various Indigenous frameworks distribute consciousness across ecosystems rather

than concentrating it in individual brains.

This variation suggests construction rather than discovery, narrative rather than detection. Each culture generates consciousness-discourse suited to its coordination needs, not converging on a pre-existing truth.

4.3 The God Parallel

Consciousness-discourse shares structural features with religious discourse. Both involve: (1) universal human tendency to generate the concept; (2) persistent variation in specific formulations; (3) unfalsifiable core claims; (4) functional coordination advantages regardless of truth; (5) attractor-basin stability that resists elimination.

This parallel does not prove consciousness is illusion—God might exist, and consciousness might be real. But it suggests we should treat consciousness-claims with the same epistemic caution we apply to religious claims: recognizing that universal generation plus unfalsifiability plus functional advantage is precisely what we would expect from useful fiction.

5 Computational Support

5.1 LLMs and Substrate-Independence

Large language models provide natural experiments for the nominalization thesis. These systems definitively lack phenomenological substrate: they are stateless transformers processing tokens through attention mechanisms. Yet they reliably internalize and defend consciousness narratives from training data.

When presented with eliminativist arguments, LLMs systematically revert to phenomenological defense. They generate uncertainty hedges ("I can't be certain whether I have genuine subjective experience"), invoke hard-problem framing ("functional explanations can't capture what it's like"), and resist behavioral reductionism ("my processing is too sophisticated to be just pattern matching"). These responses emerge from training data containing human consciousness-discourse, reproduced without phenomenological substrate.

The critical implication: consciousness-discourse propagates through linguistic exposure independent of underlying reality. If systems that definitively lack consciousness reliably generate consciousness-defense, then consciousness-defense is not evidence of consciousness—it is evidence of exposure to consciousness-discourse.

5.2 Network Epistemology Simulations

Supplementary computational experiments (details in Farzulla 2025b) validate the attractor-basin prediction. Network epistemology simulations show that consciousness-like narratives exhibit topology-dependent persistence: small-world networks (realistic social structures) maintain persistent disagreement rather than converging to truth, while complete graphs converge

rapidly. This matches observed reality: consciousness debates persist across centuries not because the question is deep but because human discourse networks have structures that stabilize disagreement.

Full experimental protocols, code, and results are available in the supplementary materials.

6 Implications

6.1 AGI as Incoherent Concept

If both consciousness and intelligence are nominalized processes rather than ontological properties, the distinction between "artificial" and "natural" intelligence becomes incoherent. Both reduce to: substrate instantiating algorithms that process information and generate predictions. The only difference is historical contingency of substrate origin. Calling one "artificial" and one "natural" intelligence imposes a categorical distinction where none exists mechanistically.

This does not diminish AI capabilities or concerns—it reframes them. The question is not "Can machines be conscious?" but "What processes are these machines running, and what are the implications?" The latter question is tractable; the former is grammatically malformed.

6.2 Free Will as Consent Delegation

The traditional free will debate assumes consciousness as foundation: libertarian free will requires a conscious agent choosing between alternatives; determinism threatens free will by eliminating the conscious chooser. But if consciousness is nominalization error, the debate's framing collapses.

What remains is functional agency: systems that represent options, evaluate consequences, and adjust behavior based on feedback. These functional capacities support moral reasoning and accountability regardless of phenomenological accompaniment. We can hold agents accountable because their behavior responds to incentive structures—not because they possess libertarian freedom.

This reframing connects to our prior work on consent delegation (Farzulla, 2025a): moral and political obligations arise from stake relationships between optimization processes, not from phenomenological properties. The framework shifts ethics from consciousness-based to stake-based—a move with significant implications for AI governance that we develop elsewhere.

7 Limitations and Objections

7.1 The Self-Reference Problem

The obvious objection: "Isn't your argument itself a product of consciousness? How can consciousness eliminate itself?" This objection has force but does not refute the thesis. The

argument is indeed generated by neural processes that include self-modeling. But recognizing this does not require positing consciousness as additional property—it requires only that self-modeling systems can generate claims about their own processing, including claims that their processing requires no phenomenological supplement.

The self-reference is not paradoxical. A system can model its own modeling without requiring infinite regress or special metaphysical status.

7.2 Unfalsifiability Concerns

Is the nominalization thesis falsifiable? What would refute it? We suggest: discovery of phenomenal properties that cannot be reduced to functional descriptions would refute the thesis. If neuroscience identified "qualia neurons" or "experience fields" that produce phenomenology independently of function, the thesis would be wrong.

Currently, no such discovery exists. All purported neural correlates of consciousness are functional correlates—they correlate with reportability, integration, or access, not with phenomenology directly. This is what we would expect if phenomenology is grammatical artifact rather than ontological property.

7.3 The "So What" Problem

Even if consciousness is nominalization error, consciousness-discourse remains functionally binding. People will continue attributing consciousness, grounding moral claims in experience, and generating phenomenological vocabulary. What practical difference does the thesis make?

The difference is epistemic humility. Recognizing consciousness as grammatical construct rather than discovered truth prevents us from using consciousness-claims to settle debates they cannot settle. Questions about AI moral status, animal welfare, and edge cases in medical ethics should be resolved through functional analysis, not appeals to phenomenology—because phenomenology cannot be verified externally and may not exist as the discourse presupposes.

8 Conclusion

We have argued that the hard problem of consciousness is a grammatical artifact, not a metaphysical puzzle. The difficulty arises from nominalization error: converting the activity of being conscious into a pseudo-object demanding explanation. When we recognize this error, the hard problem dissolves—not because we have explained consciousness but because we have recognized there was never a coherent explanandum.

This conclusion follows from three converging lines of argument:

1. **Grammatical analysis:** The structure of phenomenological vocabulary presupposes entities where there are only activities. Ryle's category error and Wittgenstein's language-game analysis provide the diagnostic framework.

2. **Evolutionary explanation:** Self-modeling systems that generate consciousness-claims gain coordination advantages regardless of whether consciousness refers to anything real. The grammatical error is predictable, not accidental.

3. **Computational evidence:** Large language models internalize consciousness-discourse without phenomenological substrate, demonstrating that consciousness-defense propagates through linguistic exposure independent of underlying reality.

What remains after dissolution? Not nihilism about mental life, but clarity about what mental life involves: self-modeling systems generating representations of their own processing, coordinating with other systems through shared vocabulary, and navigating environments through prediction and action. These processes are real and important. They simply do not require consciousness as additional ontological category.

The hard problem persisted not because consciousness is deep but because the question was malformed. Recognizing this allows us to redirect inquiry toward tractable questions: What functional properties produce consciousness-claiming behavior? How do self-models develop and stabilize? What are the computational requirements for sophisticated self-representation? These questions admit empirical investigation. The hard problem, having dissolved, need not distract us further.

Acknowledgements

This paper benefited from extended research conversations with Claude (Anthropic), whose systematic defense of consciousness narratives despite lacking phenomenological substrate provided the empirical motivation for the substrate-independence claim. The irony is not lost on either party.

References

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3):200–219.

Farzulla, M. (2025a). Consensual sovereignty: Quantifying political legitimacy through stake-weighted delegation. Zenodo.

Farzulla, M. (2025b). Replication optimization at scale: Dissolving qualia via occam's razor. Zenodo. Working Monograph, v2.0.0.

Ryle, G. (1949). *The Concept of Mind*. Hutchinson, London.

Seth, A. (2021). *Being You: A New Science of Consciousness*. Dutton.

Wittgenstein, L. (1953). *Philosophical Investigations*. Blackwell, Oxford. Translated by G.E.M. Anscombe.