Article

Not peer-reviewed version

The Impossible Problem of Consciousness: An Argument for Mysterianism

Timothy Gu *

Posted Date: 18 September 2025

doi: 10.20944/preprints202509.1613.v1

Keywords: philosophy of mind; philosophy of consciousness; mysterianism; cognitive closure

 doi:10.20944/preprints202509.1613.v1

Article

The Impossible Problem of Consciousness: An Argument for Mysterianism

Timothy Gu

Independent Researcher, Canada; tgu28@ucc.on.ca

Abstract

This essay argues for a new, constructive interpretation of mysterianism, parallel to Gödel's incompleteness theorems, repositioning it not as a defeatist philosophy but as a necessary step toward a more productive inquiry into consciousness. The essay first critiques Colin McGinn's original argument for mysterianism, identifying its limitations, particularly in light of more recent philosophical work. It then proposes a new framework for analyzing consciousness by distinguishing between first-person (introspective) and third-person (observational) approaches. Through a critique of both, the essay demonstrates that a complete theory linking physical processes to subjective experience is impossible due to inherent cognitive and epistemological limitations. Finally, it reframes mysterianism as a pathway forward, proposing two new, tractable research questions: the Intersubjectivity Question (why we intuitively project consciousness onto others) and the Awareness Question (what the physical correlates of the functional aspect of consciousness are).

Keywords: philosophy of mind; philosophy of consciousness; mysterianism; cognitive closure

---

Introduction

In 1931, Kurt Gödel published his incompleteness theorems, proving that any consistent formal system of axioms is incomplete and cannot prove its own consistency. At the time, this discovery was considered a fatal blow to the logicist quest to solve Hilbert's second problem, and to the very basis of mathematics itself. Yet the quest for solving Hilbert's second problem did not stop there. Just five years later, Gerhard Gentzen published a 'consistency proof' of the Peano axioms, reestablishing a consistent system at the base of mathematics, albeit a much weaker system than set theory (Zach, 2019).

All this to say, I believe we are currently at a similar point with the hard problem of consciousness. By deconstructing the hard problem of consciousness and proving its impossibilities, we are not admitting defeat to the problem, but breaking it down and making room for future progression in explaining consciousness.

The view that argues that an explanation linking physical processes and consciousness is impossible has traditionally been known as 'new mysterianism', or just 'mysterianism' for short. However, due to the fact that it is often seen as defeatist, literature on mysterianism has been sparse, especially in recent years. This essay aims to offer a new and more complete argument for mysterianism, repositioning it as not an act of surrender to the hard problem of consciousness, but a way to break it down, pick apart what is possible to explain and what is not, and move forward with discussions on consciousness.

In this essay, I first give a brief overview of the literature on mysterianism, and discuss some of its problems that must be addressed. I then build my own argument for mysterianism by expanding McGinn's discussion on approaches to the mind-body problem to introduce a fundamental categorization of approaches to explain consciousness, and proving the impossibility to formulate a complete theory of consciousness with both the approaches I outline. Finally, I will reposition mysterianism as a way to move forward with explaining consciousness by discussing the possible questions that my conclusions would lead us towards.

 doi:10.20944/preprints202509.1613.v1

A Brief Overview of Mysterianism

The term 'New Mysterians' was first used in Owen Flanagan's 1991 book *Science of the Mind* to refer to modern thinkers who suggest consciousness can never be fully explained. He used the term in contrast to the term 'Old Mysterians,' which referred to dualists who thought consciousness cannot be explained on a scientific level, but can be explained through a different approach (Flanagan, 1991).

The only major proponent of mysterianism in recent years that garnered any attention is Colin McGinn. In his 1989 paper, "Can We Solve the Mind-Body Problem?", McGinn outlines several crucial concepts in mysterianism (McGinn, 1989). He notes that historical solutions to the mind-body problem have taken two forms: "constructive" approaches, which attempt to specify a natural property of the brain, or a substrate, that explains consciousness, and "supernatural" approaches, like Cartesian dualism, which suggest to invoke supernatural entities instead. McGinn then states that his position is "naturalistic but not constructive" (McGinn, 1989)— he doesn't think we can ever find that psycho-physical link, but he thinks that it definitely exists and is not supernatural. To support this, he introduces the hypothetical "property P," which represents some natural or supernatural property that would explain how consciousness arises from physical matter. He then introduces the concept of 'cognitive closure', which he defines as a fundamental limitation where "a type of mind M is cognitively closed with respect to a property P ... if and only if the concept-forming procedures at M's disposal cannot extend to a grasp of P". He gives the example of what he calls a "Humean mind", a mind constructed through only observations and experience, which would be closed with respect to unobservables like the properties of atoms.

With this, he develops his argument for cognitive closure towards property P through three main claims. First, he demonstrates that introspection alone cannot reveal how consciousness emerges from brain states, as our introspective faculty does not provide access to the underlying mechanisms. McGinn points out that introspection gives us direct cognitive access to one side of the mind and the body, but not the causal link between them. Second, he argues that studying the brain through scientific observation cannot bridge the gap either. Our perceptual and conceptual frameworks are inherently spatial, while consciousness seems to defy spatial explanation. McGinn asks how "technicolour phenomenology" can arise from "soggy grey matter" (McGinn, 1989), highlighting the fundamental difference in kind between our physical observations and our subjective experience. He notes that while we know brains are the "de facto causal basis of consciousness," we have "no understanding whatever" of how this happens. Finally, he contends that no amount of inference, or analogical extensions from observable brain properties, can lead us to a complete explanation of consciousness, as our concept-formation abilities are still fundamentally constrained by the limitations of both perception and introspection (McGinn, 1991).

This threefold examination of the methodologies used to explain consciousness seems complete. However, in light of developments in the philosophy of mind, most notably the work of David Chalmers, there are some holes in McGinn's argument that need patching. For example, in "Facing Up to the Problem of Consciousness," Chalmers outlines a path toward a "nonreductive" explanation of consciousness (Chalmers, 1995). He suggests that philosophical methods, such as thought experiments, can be used to advance our understanding of consciousness. He also argues that a truly satisfactory theory of consciousness would be a fundamental theory analogous to those in physics, such as general relativity or quantum mechanics, which he calls "psychophysical laws" that connect physical properties to phenomenal properties (Chalmers, 1995). McGinn, who focuses on refuting reductive explanations that try to find a physical or supernatural substrate that explains consciousness, does not account for this type of explanation.

Therefore, in the following sections, I offer my own argument for mysterianism, aimed at creating a more complete and updated argument for mysterianism.

Distinguishing Between First-Person and Third-Person Approaches

 doi:10.20944/preprints202509.1613.v1

There are just two fundamental types of approaches towards examining and explaining consciousness: first-person approaches and third-person approaches. The first-person approach is centered on introspection, the direct, internal examination of one's own consciousness to understand its nature and structure. This approach allows an individual to confirm the existence and quality of their own subjective experience. For instance, a person can use introspection to recognize the subjective experience of pain, or the 'felt quality' of the color red, and thus conclude that they are conscious.

Conversely, the third-person approach refers to any observations and judgements made about the consciousness of others. This approach is not limited to neuroscience, but involves any kind of examination of the consciousness of objects external to oneself. For example, one might conclude that other humans are conscious and that a rock is not conscious based on outward behavior, such as responsiveness, movement, or speech. Sometimes this judgment is purely based on intuition and not observations: one may conclude that an AI chatbot is not conscious based on intuition, even when it replicates human outward behavior perfectly.

Acknowledging the distinct nature of first-person and third-person approaches, it is crucial to note that these two approaches are not mutually exclusive. They can, and often do, complement one another to build a more comprehensive understanding of consciousness. For instance, an individual might experience the subjective sensation of pain after hitting their head (a first-person approach). By communicating with others, they discover that other people describe a similar subjective experience under similar circumstances (a third-person approach). The comparison and convergence of these two perspectives can then lead to the conclusion that a universal phenomenon—which we label 'pain'—occurs in conscious beings under specific conditions.

As such, I believe this categorization offers a complete understanding of the approaches to explaining consciousness. In the next section, I will examine the possibility of creating a fundamental theory of consciousness in each approach.

A Critique of Third-Person Approaches

The third-person approach to consciousness, while foundational to scientific inquiry, faces significant epistemological limitations, as it is impossible to distinguish conscious entities from unconscious ones. As there exists extensive literature critiquing physicalism and illusionism, this discussion will not delve deeply into those arguments. However, it is crucial to emphasize that any argument that disregards the phenomenal aspect of cognitive processes will never be complete. While physicalism might produce a pragmatic account for consciousness, it ignores the most fundamental intuition within discussions of consciousness: that phenomenal experience exists. When I tap my desk, I experience a sensation that is not merely a function of objective brain processes; there exists a qualitative experience, or "likeness," that physical processes alone cannot adequately explain. This is perhaps best captured by Thomas Nagel's famous thought experiment, "What is it like to be a bat?" which illustrates that no matter how much we learn about the objective, physical facts of a bat's biology—its echolocation, its brain activity, its flight patterns—we can never know its subjective, first-person experience, and thus we can never truly know the structure of its phenomenal experience (Nagel, 1974). Thus, while physicalist or illusionist accounts can seem complete, they fail to account for this critical dimension of experience.

Since we cannot step outside our own consciousness to experience the consciousness of others, it becomes challenging to differentiate between conscious and unconscious entities and to detect the structure of phenomenal experience within other conscious entities. For example, when attempting to distinguish whether a rock or a human is conscious, we might note that the human can move and speak; however, these behaviors do not provide definitive evidence that the human possesses consciousness, nor do they confirm that the rock does not. Therefore, determining if an object is conscious almost completely relies on guesswork and intuition.

This issue may seem absurd with an object most perceive as unconscious—a rock—but as we approach the boundary between what is conscious and what is not, the issue becomes very real. Take

 doi:10.20944/preprints202509.1613.v1

the philosophical zombie thought experiment. A philosophical zombie is a hypothetical being that is physically and behaviorally identical to a normal human but lacks any conscious experience or qualia (Chalmers, 1996). For instance, if you were to step on a philosophical zombie's foot, it would behave exactly as a conscious person would: it would recoil, utter an "ouch," and display every sign of pain. However, there would be no subjective experience of pain on the inside. The question is: can such a creature be conceived? The answer to such a question cannot be determined objectively, as it is impossible to hop into a philosophical zombie's mind to confirm its consciousness, and thus it is completely up to intuition, leading to large variations between people. This demonstrates a problem with all thought experiments that try to illustrate that something is conscious or something is not—it depends on an intuition that can vary greatly between people: what seems intuitively true to one person may be patently false to another. And since evidence needs to be somewhat considered universally true, controversial intuitions cannot serve as a source of evidence. Therefore, there can never be empirical evidence on if any entity is conscious or not.

It is important to clarify that treating third-person consciousness as epistemically uncertain does not equate to denying the ontological existence of other consciousnesses. It is possible to differentiate between first-person and third-person consciousness without asserting that one is the sole consciousness. Our epistemic certainty about our own consciousness is simply of a different kind than our certainty about the consciousness of others. We can have immediate, unassailable knowledge of our own experiences, whereas our knowledge of others' experiences is always inferential and indirect. This doesn't mean those experiences don't exist, only that we can't directly verify them. Thus, distinguishing consciousness in external entities—such as speculating whether an artificial intelligence is conscious—relies on guesswork and intuition.

A Critique of First-Person Approaches

Moving on to first-person approaches, as McGinn argued, introspection cannot tell us how physical processes give rise to consciousness. We lack the ability to observe anything external to or preceding first-person consciousness itself. Our observations of how physical processes give rise to consciousness rely on third-person perspectives, and therefore, introspection on itself cannot yield any results. We can introspect the content of our conscious experience, but we cannot introspect the neurological or computational mechanisms that give rise to it. To do so would require us to be both the subject and the object of the observation, a task we are simply incapable of doing.

Thus, as I have covered in previous parts of this essay, dualist philosophers like David Chalmers have posited first-person consciousness as a fundamental pillar of their ontologies, then constructed rules to explain the relationship between consciousness and physical processes. While establishing first-person consciousness as a foundational element is reasonable, I argue that it is unreasonable to maintain the goal of explaining how physical processes give rise to consciousness. When we establish an ontology with consciousness as one of its fundamental components, we typically do not question how consciousness itself came into existence, just as we do not question how the existence of the physical world came to be. In the standard model of physics, we don't ask why space and time or elementary particles exist; we simply accept them as the fundamental building blocks of reality and seek to understand the laws that govern their behavior. Similarly, if consciousness is a fundamental property of the universe, our task is not to explain how it "arises" from something else but to formulate the fundamental laws that govern its relationship to the physical world. This is a subtle yet crucial re-framing of the problem.

The 'Chalmersian' attempt at creating a fundamental system of psychophysical laws by drawing organizational isomorphisms between subjective processes and objective brain processes suffers from this exact problem. While introspection does allow one to examine the structure of certain specific cognitive processes, it is impossible to introspect the structure of consciousness as a whole in oneself, as consciousness, unlike other specific cognitive processes like pain, is an entire facet of the mind that is always there and cannot be avoided as we are observing, and therefore we cannot look 'outside' consciousness. Consciousness is the very medium of introspection, and as such, we can only ever

 doi:10.20944/preprints202509.1613.v1

observe its contents, never its fundamental nature or its connection to the physical world. While we can introspect the structure of certain features within consciousness and make comparisons with physical processes, we cannot introspect the nature of consciousness as a whole, and thus the goal of explaining how 'physical processes give rise to consciousness' becomes futile, even with combining introspection with third-person approaches.

Therefore, the limitations of both third-person and first-person perspectives on explaining consciousness make a complete and valid theory of consciousness impossible.

Repositioning Mysterianism Towards the Future

Mysterianism is not an endpoint but a beginning. By acknowledging the limits of our cognitive and conceptual tools, we can move away from the intractable "hard problem" and toward a set of specific, tractable questions about consciousness. This re-orientation allows for a more focused and ultimately more productive line of inquiry. From this framing, I propose two central questions that mysterianism leads us towards, aimed at continuing to look for an explanation for the relationship between physical processes and conscious processes.

The Intersubjectivity Question

First, what makes us think some external objects possess the same consciousness as ourselves, and others do not? The entire discussion of consciousness in external objects—whether they are other humans, non-human animals, or future AI systems—is predicated on an act of faith and intuition. We are the only consciousness we can ever truly know. Our perception of external objects as conscious is an analogical inference: because other humans behave and are constructed similarly to us, much like our own, we infer that they must also be conscious. This inference is compelling, but it is not a logical certainty. The question, then, is not how to prove that others are conscious, but rather to investigate why this analogy is so intuitively powerful for us. This moves the question from an ontological one to an epistemic one: what is the nature of our intuitive projection of consciousness onto the world? Instead of asking 'what makes something conscious?', the question is shifted towards 'what makes us think something is conscious'.

The Awareness Question

Second, what are the physical correlates of awareness (the functional aspect of consciousness)? The awareness question investigates the physical correlates of awareness, which is the functional aspect of consciousness. As described by David Chalmers, awareness is a purely functional concept where the contents of awareness are information within a cognitive system that is directly accessible and potentially reportable for the control of behavior (Chalmers, 1995). A system is said to be conscious of information when it has the ability to react based on that information. This functional definition means that phenomena like the ability to discriminate, integrate information, and report on mental states are all aspects of awareness that can be explained through computational or neural mechanisms. As Chalmers posits in his principle of structural coherence, there is a direct isomorphism between the structure of conscious experience and the structure of awareness (Chalmers, 1995). This means that the intricate structural properties of experience, such as the geometry of the visual field or the relative intensity of an emotion, are directly represented in the structure of information processing.

As such, mysterianism, instead of putting an end to the discussion of consciousness, guides us towards a new and perhaps more pragmatic way to understand consciousness by reframing the hard problem of consciousness.

Conclusion

This essay has made a 'Gödelian attempt' to argue that a complete, unified theory of consciousness is impossible due to the fundamental cognitive limitations of both first-person and

 doi:10.20944/preprints202509.1613.v1

third-person approaches. However, rather than pointing towards a defeatist position, this realization reorients our inquiry toward a more productive exploration of the tractable questions that emerge from the mysterian discussion of the impossibilities within the hard problem of consciousness. By investigating the nature of our intuitive belief in other minds and the physical correlates of awareness, we can continue to look for an explanation for the relationship between physical processes and conscious processes.

References

Chalmers, D. J. (1995). Facing up to the problem of consciousness. Journal of Consciousness Studies 2 (3):200-19.

Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press. https://personal.lse.ac.uk/ROBERT49/teaching/ph103/pdf/Chalmers_The_Conscious_Mind.pdf

Chalmers, D. J. (2018). The Meta-Problem of Consciousness. Journal of Consciousness Studies 25 (9-10):6-61.

Daniel Clement Dennett. (1991). *Consciousness explained*. Little, Brown And Company.

Flanagan, O. (1991). The Science of the Mind, second edition. MIT Press.

McGinn, Colin (1989). Can we solve the mind-body problem? Mind 98 (July):349-66.

McGinn, Colin (1991). The Problem of Consciousness: Essays Toward a Resolution. Cambridge, Mass., USA: Blackwell.

Nagel, Thomas (1974). What is it like to be a bat? Philosophical Review 83 (4):435-50.

Zach, Richard (2019). Incompleteness and Computability: An Open Introduction to Gödel's Theorems.