

The Eye of The Mirage: On the Elusive Nature of Consciousness

Sepehrdad Rahimian¹

¹ Department of Theoretical Philosophy, Research University of Groningen, Netherlands

* Correspondence:

Sepehrdad Rahimian

s.rahimian@rug.nl

Abstract

Is there a distilled, core, phenomenal feel that sits at the heart of conscious experience? In a usual debate about consciousness, there are often two sides involved: the physicalist, who reminds us that many mysteries of the past have now been shown to have been entirely misguided notions as our scientific understanding of the world has advanced, and the anti-physicalist, who emphasizes that consciousness stands apart since our very understanding of science is filtered through the senses, making consciousness the very existence one cannot possibly doubt. Clearly, regarding the topic of consciousness, even the explanandum is heavily debated, and perhaps finding a neutral explanandum is practically impossible, even if one believes in a perfect separation of facts and theory. By pointing out the limitations of both physicalist and non-physicalist ontologies, this paper explores the possibility of a neutral explanandum and argues even if one exists, it has an elusive nature.

Keywords

Consciousness, Antirealism, Qualia, Epistemic Barrier, Physicalism

1. Introduction

“The only principle that does not inhibit progress is: anything goes.”

- Paul Feyerabend (2020), “Against Method”, p.14

How does one come to learn something about consciousness? Does a color scientist, who has been kept in a black and white room her entire life, learn something new about color the first time she sees a red tomato? What if she already possessed knowledge of all the physical facts about vision, color, neuroscience, etc.? I believe the most overlooked aspect of Frank Jackson’s famous thought experiment (Jackson, 1982) is not the setup or the question itself, but rather, what motivates that question and similar ones in the first place. When contemplating any thought experiment, philosophers could typically take one of two routes: The first route is to

directly address the question itself (i.e. *fine-grained* philosophy). For instance, with respect to the Mary thought experiment, those already convinced that consciousness cannot possibly be physical would give a definitive "yes" as an answer. Others may point out the strength of the premise, emphasizing that since consciousness is physical and Mary knows *all* the physical facts about it, she learns nothing new by seeing color for the first time. Fine-grained philosophy allows us to *zoom-in* on one clearly defined issue and to explore it with as much precision as possible. This precision however, sometimes comes at a cost of missing out on a bigger picture. This is where the second route, or the *coarse-grained* philosophy comes into play. Coarse-grained philosophy allows us to *zoom-out* from the matter at hand through *relaxing* the semantics and definitions with the aim of addressing more general patterns. For instance, one could explore the general motivation behind the "Mary-type" thought experiments and quickly find out that they all seem to argue, one way or another, that there is "something" about subjective experience that resists scientific explanation.

Clearly, both modes of analysis (i.e. fine-grained and coarse-grained¹) have their importance, but it is the latter one that I intend to focus on in this paper. I believe much has been said on the subtle distinctions between various consciousness-related constructs and that this has somewhat diverted our attention from discussing the "prime question" without getting side tracked, philosophically. One way to formulate the prime question is the following:

"There is an aspect of my existence (or my experience) that feels like something. What is the nature of this *aspect*?"²

This paper aims to explore that aspect of consciousness which the prime question is about and to point out an epistemic veil that bounds our access, irrespective of whether or not something lies beyond that veil. While I will touch upon various controversial topics in the philosophy of mind, my primary focus shall remain on laying the groundwork to explore the possibility and nature of a neutral explanandum using a more coarse-grained approach. Hence, I shall address these topics mainly at a surface level to the extent that they can help us explore the prime question.

In section 2, I briefly review the idea that modern science from the beginning may have been set up to fail at explaining consciousness. Section 3 considers the possibility of zombies and examines whether they are genuinely incompatible with physicalism. I emphasize how arguments both in support and in opposition to zombies can exhibit elements of circular reasoning. By acknowledging this circularity, I suggest refraining from outright a priori dismissal

¹ This distinction is in part inspired by Dewhurst (2021).

² The prime question is similar to the hard-problem (Chalmers, 1996) for most parts, it simply emphasizes more on exploring the nature of consciousness to whatever end it leads and less on the relationship between mind and matter. The prime question is agnostic about this relationship and still holds whether mind and matter are of the same nature, or of different natures. It is possible to say the same about the hard-problem, but that would be open to interpretation.

of any ontology, regardless of how unconventional that view may appear to the reader. In sections 4 and 5 I strive to distill the explanandum as it relates to the prime question, setting aside unnecessary semantic and theoretical distractions to uncover the heart of the matter. An analogy here is akin to establishing the existence of the earth before debating its roundness. Section 6 ends by considering the possibility that the first-person and the third-person perspectives could be complementary to one another in understanding consciousness without having to force ideas of one domain to the other.

2. Galileo, Goff, and philosophical zombies

Let us now turn our attention to Jackson's "Fred" (Jackson, 1982), who has received comparatively less attention than "Mary" but may better assist us in distilling our explanandum.³ In the thought experiment, Fred possesses superior color vision compared to the majority of people. For instance, where we see only red tomatoes, Fred sees *red1* and *red2*. He insists that the difference between the two reds is not a matter of shades of the same color, but that they are as different as yellow and blue. An examination of Fred's visual system reveals that he can indeed discriminate between two groups of wavelengths in the red spectrum, similar to how an average person distinguishes between yellow and blue. However, as Jackson argues, it appears that no amount of physical inspection can reveal the nature of Fred's subjective experience when he looks at red1 and red2. In other words, what is it like for Fred, seeing tomatoes that appear the same color to us? What is Fred seeing?

In his book "Galileo's Error" Philip Goff (2019) makes the case that modern science has been set up since Galileo to exclude sensory qualities and to only include those properties about the world that can be described mathematically, such as shape, size, length, etc. In accordance with Goff's observation, Galileo in fact thought of sensory qualities to reside only in the *sensitive body*⁴:

...I say that upon conceiving of a material or corporeal substance, I immediately feel the need to conceive simultaneously that it is bounded and has this or that shape; that it is in this place or that at any given time; that it moves or stays still; that it does or does not touch another body; and that it is one, few, or many. I cannot separate it from these conditions by any stretch of my imagination. But that it must be white or red, bitter or sweet, noisy or silent, of sweet or foul odor, my mind feels no compulsion to understand as necessary accompaniments. Indeed, without the senses to guide us, reason or imagination alone would perhaps never arrive at such qualities. For that reason I think that tastes, odors, colors, and so forth are no more than mere names [*puri nomi*] so far as pertains to the subject wherein they appear to reside, and that they have their

³ I am grateful to an audience member who raised this point during Q&A of a debate between Sean Carroll and Philip Goff.

⁴ As cited in Piccolino and Wade (2013). Elsewhere, he also mentions these qualities as residing in the *soul*, but Galileo's notion of the soul differs from today's Cartesian understanding. The provided paragraph adequately represents the essence of Galileo's view on sense data.

habitation only in the sensitive body. Thus, if the living creature [*l'animal vivente*] were removed, all these qualities would be removed and annihilated... (Saggiatore, pp. 196–197; transl. pp. 309–310)

In essence, what Galileo seems to have thought is that on one hand there are objective properties that are independent of our existence (and are measurable and worth theoretical consideration) and on the other, there are subjective qualities that only exist for the subject itself. Therefore, if the subject were to be removed, these qualities would also cease to exist⁵.

Here we can highlight one of Goff's central claims that is of utmost importance in studying consciousness. Although modern science has been vastly successful in explaining and predicting the world around us and that mathematics has played a crucial role in this achievement, it does not follow that whatever science (i.e. physicalism) cannot touch upon is *unreal*. Goff then provides the zombie argument (Chalmers, 1996a) to serve as an objection to physicalism:

1. If materialism is true, then feelings are identical with brain states.
2. If feelings are identical with brain states, then it is not logically possible for feelings to exist without brain states, or vice versa. (This follows from the identity principle.)
3. If zombies are logically possible, then it is logically possible for brain states to exist without feelings.
4. Therefore, if zombies are logically possible, materialism is false.
5. Zombies are logically possible...
6. Therefore, materialism is false (Goff, 2019, pp. 87-88).⁶

Goff clarifies his use of “logically possible” by pointing out that neither “flying pigs” nor “round squares” exist, but the former has an advantage over the latter since it is possible to imagine how evolutionary processes could have produced pigs that can fly, but the existence of round squares is simply a contradiction. Thus, the former is a logical possibility, whereas the latter is not. In short, the argument can be summarized as the following: If zombies are possible, then physicalism is false, zombies are possible therefore physicalism is false.

3. Do zombies really refute physicalism?

⁵ Galileo's view on the “senses” is explored in depth in Piccolino and Wade (2013). Also, a novel classification of Qualia into Galilean and non-Galilean is explored in great detail in Sundström (2014).

⁶ “Galileo's error” is a book targeting mainly a general audience as opposed to professional philosophers and Goff has simplified some concepts to improve readability. However, at least in the context of this paper, I find this simplified and easy-to-follow language more useful. My critique of the zombie argument would remain the same either way.

A standard response to the zombie argument has been to point out that the mere “conceivability” of X does not entail the “possibility” of X (Kirk, 2023).⁷ However, I believe that responses of this sort, while potentially valid, are as unhelpful as the zombie argument itself is, when used solely as a response to physicalism. This is mainly because they preach to the already converted and they are unlikely to change anyone’s mind about the issue. Therefore, while exploring the relationship between conceivability and possibility remains an interesting topic in its own right, here I shall proceed under the assumption that zombies are indeed possible.

The main challenge with Goff’s argument in my view is in fact bringing the identity principle into the equation, which I believe is unwarranted. More specifically, the issue is using an argument that depends on an absolute truth claim regarding the mind-body identity (and in his case, lack-there-of) for soundness. Goff’s use of identity principle is the standard metaphysical one; that is, if A is identical to B, then A must be identical to B in any possible world where A and B exist. The problem is that such strong metaphysical relationship between consciousness and brain states is itself a major point of controversy and ought not to be taken for granted in either direction, for that if the truth of the matter was so plainly obvious, there would not have been many different sides to consciousness debates. For his argument to hold, one must axiomatically assume already that brain states are not the same as conscious states (to provide a basis for conceivability and then possibility of zombies) and that this is true of every possible universe including ours which is essentially what is implied by the argument and hence, the argument is circular. If we can already grant that physical states (e.g. brain states) are not the same as conscious states, then we have already rejected physicalism in the first place and zombies are the product of this belief; they do not lend additional support to it.

The circularity of course cuts both ways, since it is evident how another argument can be constructed, holding the assumption that brain states *are* identical to conscious states, and then concluding that zombies would necessarily have every bit as consciousness as we do and thus, their existence is a logical contradiction⁸. Then, the argument could even continue down this road to rule out the possibility of more “exotic” ontologies (e.g. panpsychism) based on this supposed inconceivability of zombies.

Regardless of whether one accepts the possibility of zombies or that one views it akin to considering the possibility of liquid H₂O without it being water, I am not convinced that zombies automatically refute physicalism either. While zombies are inconsistent perhaps with a radical brand of physicalism that asserts everything that is real in every possible universe is

⁷ There are various objections to zombies including verificationist objections, epistemic contact, and more. These objections are further explored in Kirk, 2023.

⁸ In other words, my issue is not with the identity principle per se, but rather, my claim is that for identity principle to be relevant in consciousness debates and to function properly in that context, it needs to be plugged in with a particular *datum* (e.g. consciousness *is* identical to brain states, or that it is *not* identical to brain states). Since I argue we are in no position to make an absolute claim regarding this datum in either direction, we cannot utilize the identity principle to anyone’s rescue.

always physical and only physical, the notion that *our* world is physical, even if contingently so, remains unaffected by the possibility of zombies. For instance, zombies and *real* conscious people could exist in different worlds and ours could just happen to be one that is compatible with physicalism. Thus, zombies could still be *possible*, but the laws of “our” world would not allow for them to exist *here*.⁹ It could be that in our world, it is not possible for right kinds of brain states to exist without their corresponding feelings and therefore, all that is required to explain consciousness is physicalism. At the same time and in another world, to explain consciousness one would have to go above and beyond physicalism.¹⁰

Adopting this *open-season* approach towards a priori ontological assertions allows for the possibility of *illusionism* as much it does for zombies. According to illusionists, we do not experience phenomenal consciousness, we merely think that we do, or in other words, conscious experience lacks phenomenal properties (Frankish, 2016). Although many may consider this view to be unsatisfying or even *crazy* (Frances, 2008), an armchair refutation of illusionism is not as easy as it is tempting (Rahimian, 2022). As Kammerer aptly notes:

The main reason illusionism about consciousness tends to be rejected is that it seems crazy; but it is in turn quite plausible that the main reason it seems crazy is precisely that we encounter deep difficulties when we try to represent to ourselves that phenomenal consciousness does not exist even though it seems to exist (Kammerer, 2021, p.851).

The possibility of illusionism of course does not automatically render physicalism as the ultimate metaphysical description of the universe either. There are serious objections regarding physicalism that are explored in, for example, Stoljar (2024) and Montero (2001; 2012). However, for the purposes of this paper, I shall continue as an “as-if” physicalist. In his seminal essay on physicalism, Stoljar (2024) raises the question of what truly distinguishes believing physicalism from merely engaging in the pretense of accepting it. One potential response is that the distinction lies in the *degree of commitment*. Considering the remarkable success of physical sciences to explain a myriad of phenomena, I am inclined to *bet* that upon confronting a new phenomenon, the tools to comprehend it will not necessitate expanding our physicalist toolkit with additional entities, specially through ways that are incompatible with the general epistemic consensus among physicists. Physicalism, then, can be seen merely as a set of practical assumptions that have proven useful so far and have led to great success in many areas of inquiry, although arguably not all – as evidenced by the ongoing debate surrounding

⁹ It is important to note that physicalists need not and do not reject the possibility of non-physical worlds. What physicalists usually cannot accept, is the claim that there could be a world that is physically identical to ours, except with no consciousness in it.

¹⁰ Admittedly, these two notions of consciousness are not identical in the metaphysical sense, but they could be considered identical epistemically. Even if there is an ultimate truth regarding this matter, given our limited vantage point we cannot be certain as to whether our notion of consciousness is epistemic or metaphysical. The argument I make later on, will help to clarify this further. I remain skeptical that any strong metaphysical claim regarding the identity of conscious states and brain states is justifiable beyond mere assertion.

consciousness. To claim that these set of assumptions are necessary truths about the world, or will always remain true, is a leap of faith I am not ready to take. For adherents of physicalism, there is comfort in identifying the *known knowns*. As an “as-if” physicalist, I cannot help but to wonder how *unknown unknowns* could reshape our understanding of what we currently think to be known knowns.

Essentially, I believe that neither embracing the possibility of zombies requires rejecting physicalism, nor does the success of science in its intended domains entail that the metaphysical claims of physicalist doctrine are necessarily true. For a topic as complex as consciousness where the range of disagreement varies from believing in the hard-problem (Chalmers, 1996b) or in phenomenal consciousness (Block, 2002) to rejecting the hard problem entirely (Dennett, 1993), it may prove fruitful to avoid (or post-pone) dismissing opposing ideas simply on the basis of metaphysics.

4. Distilling qualia

Let us now explore, without a priori commitments to physicalism, whether there exists an aspect of consciousness where we have reasons to suspect that science falls short in capturing. Philosophers often use qualia as such examples. There are many different formulations of qualia and different ways to categorize them, but certain properties such as being intrinsic, ineffable, direct and private seem to recur in discussing qualia (Dennett, 1988; Tye, 2021). By keeping our prime question in mind, we can already see that this formulation carries too much *theory* with itself. There is a difference between a coarse-grained, neutral (or rather, as neutral as possible) explanandum that we try to get to, and a *quale*.

In his classic paper "Quining Qualia" Dennett (1988) challenges the theoretical attributes often ascribed to qualia through a series of thought experiments and argues that these attributes are not as robust as their proponents claim them to be. One of his most relevant thought experiments to this discussion is the “osprey cry”:

...I have never heard the cry of an osprey, even in a recording, but know roughly, from reading my bird books, what to listen for: "a series of short, sharp, cheeping whistles, *cheep, cheep* or *chewk chewk*, etc; sounds annoyed." (Peterson, 1947) ...The verbal description gives me a partial confinement of the logical space of possible bird cries. On its basis I can rule out many bird calls I have heard or might hear, but there is still a broad range of discriminable-by-me possibilities within which the actuality lies hidden from me like a needle in a haystack. (Dennett, 1988)

As Dennett illustrates, a bird book can indeed assist us in the process of identifying an osprey through its sound, or at least increase our chances of distinguishing it from other birds. When combined with additional information, such as visual appearance, we can gain further confidence in our identifications. Dennett readily acknowledges that upon actually hearing an osprey's cry for the first time, he undergoes a qualitative state that appears to involve a "particular mental complex of intrinsic, ineffable qualia". Sadly, this is where Dennett leaves us,

gradually shifting the direction of his arguments away from a thorough analysis of the very state he had just so masterfully isolated and goes on to argue why the importance of this matter is overstated. He commits what I call the *illusionist sin* by dismissing the core aspect of consciousness while dedicating most of his time debunking the theoretical claims of his opposition rather than further clarifying the matter. Here is an exaggerated formulation of the illusionist sin:

“The statement XYZ about consciousness is false and inflated since Z is clearly incorrect for this and that reason. As far as X is concerned, it is tempting to think that there is an X but if you think a bit harder, XY is wrong because Y is also overrated for such and such reason so XY can’t be true either. In conclusion, XYZ is an illusion...”¹¹

Following Dennett’s footsteps, Keith Frankish argues against a theoretically confused construct that he refers to as “diet qualia”. Frankish (2012) defines “classic qualia” as the ineffable, intrinsic properties of subjective experience, and also “zero qualia” as those properties that dispose us to merely judge that our experience has classic qualia. He then, through a series of arguments, dismantles the notion of “diet qualia”, which is supposed to be *weaker* than classic qualia but *stronger* than zero qualia. He writes:

In short, I understand what classic qualia are, and I understand what zero qualia are, but I do not understand what diet qualia are; I suspect the concept has no distinctive content (Frankish, 2012, p.669).

One of the aims of Frankish’s essay is to argue against a *neutral* explanandum in understanding qualia and against the notion that if one digs deep enough into classic qualia, one can reach a certain core phenomenon where the classic qualia are about. In other words, the answer to my prime question in his view would be a version of diet qualia. I believe Frankish successfully shows that at least certain uses of diet qualia are indeed confused, and that they either collapse into classic qualia, or zero qualia.¹² However, I believe there is still something left that is not considered by Frankish too seriously.

The main approach that Frankish has taken in his essay, is to argue in which existing “theoretical box” does consciousness fit the best, and to demand precision and clarity from any general notion of phenomenal experience that defies these boxes. But what about a general, somewhat unclear notion of a core phenomenal experience? Could we really claim that if an explanandum is not *precise*, it is therefore unworthy of further exploration? This is one example where I believe that emphasizing clear definitions and fine-grained semantics risks overlooking

¹¹ And it is at this point where I experience the sharp quality of a cry inside my head, perhaps not unlike one from Dennett’s osprey: “BUT WHAT ABOUT X?!”

¹² In fact, my conclusion—that even if there is a neutral, diet version of qualia, it has an elusive nature—is in agreement with what Frankish suspects the concept to be. However, I believe we arrive at this conclusion through different routes and take the argument in slightly different directions. Frankish ultimately is an illusionist, and although I take illusionist arguments very seriously, I maintain a more agnostic position regarding phenomenal consciousness and in general, the nature of consciousness.

the broader, more general construct that we could otherwise be considering. If anything, I would prefer the explanandum to remain somewhat ambiguous rather than being first hindered by an excessive focus on details, only to be later dismissed as unclear or confused. Therefore, if I am correct, there is still an aspect of consciousness that is not properly “Quined”, and it is as general as it gets¹³ : an *ineffable* character of experience.

Here I am using the term ineffable in a strong sense. If something is ineffable, it cannot be adequately expressed in words or through mathematical formulations and symbolism. We can perhaps point to it and scratch the surface by employing analogies and examples, with the hope that others, whom we assume also experience this quality, form some understanding of what we are attempting to refer to.

5. The neutral explanandum

The ineffable character of experience, in my view, is as neutral as we can make the explanandum to be. Those who subscribe to qualia theories are likely to endorse it and may add to it other properties such as directness, privateness, etc. Some might argue further that these properties cannot be physical. On the other hand, physicalist opponents of qualia theories can still acknowledge that subjective experience is ineffable. They may emphasize, however, that our understanding of this ineffability is heavily mistaken and that there is nothing "magical" going on. This, I believe, is the point where everyone can agree that there is an "earth," so to speak, allowing for discussions about its shape.

The question then becomes: “Given the ineffable character of my experience, if a scientific paradigm sufficiently captures the reasons why the dynamic interplay among certain physical processes allows me to feel that I am undergoing that experience, has the paradigm captured everything there is to be explained?”¹⁴

At first glance, it does seem that something is left out — something I have been attempting to pinpoint which has to do with why does this *feel*, feels the way it does, rather than feeling like anything else (including nothing). The emphasis here is on unraveling the “*why*”, and not the “*how*”.

Let us designate this feel as X. What, then, is X? One answer is that X represents what must be added to a philosophical zombie to render its (illusory) subjective experience veridical. However, this response takes us right back to where sometimes philosophers assume a hidden premise first and circle back to it later on, ad infinitum. At the same time, this answer is useful in one sense: if there is any non-physical aspect to consciousness at all, it likely pertains to X.

5.1 What is the nature of X?

¹³ And for the purposes of the argument that I am going to make, this generality is a point of strength.

¹⁴ There are good reasons to think that even if one believes in physicalism, a clear-cut case of complete reductionism may prove impossible due to sheer size of the complexity barrier we will face in breaking down the entire functionality of a complex system such as the human brain (Herzog, Doerig, and Sachse, 2023).

Since we are entertaining that X might be non-physical, let us momentarily set aside science. Can philosophy shed light on the nature of X? Can *mysticism*? Do we have any compelling reasons to believe anything to be true about X, aside from simply asserting its *obvious* truthfulness ad nauseam? Consider the following thought experiment:

Suppose there are two worlds each with its own earth. In world one, people are *really* conscious. Their experience includes X and when they say “I feel pain”, there really is something like for them to feel pain. In world two, everyone is a zombie. Their experience only gives the appearance that it has X, but in reality, it is empty. World two is an *X-less* world and the reports of pain are non-veridical. If a group of interworldly Martians with a highly advanced understanding of neuroscience were to analyze samples of individuals from each world, given that X is non-physical, they would be unable to distinguish between the two groups. The Martians would study their behavior, scan their brains, even form functionalist theories about their cognitive states if they felt the need to summarize their calculations and use as many “centers of gravity” as they have to so that the earthlings could also understand how their brains work. The Martians would develop a theory of mind and yet, they would entirely overlook X. How could they possibly know about X, the nonphysical, ineffable quality of experience? From the Martian perspective, everyone would be considered a zombie, leading them to be fundamentally mistaken about an entire planet of people. The way we have set X up makes it unsurprising that a third-person assessment falls short of capturing it. But what about a *first-person* assessment?

Suppose I told you that you either belong to world one or world two. Can *you* discern which world you belong to? You know what it is like for you to experience the redness of red, the searing bite of pain, or the oak-kissed scent of fine whiskey; you know your experience is imbued with X. However, so does the zombie you. The zombie you would be under the illusion that it has all the similar qualities, including X, but in fact it does not; it has every quality except X. You can experience how it feels to feel pain, but you cannot know whether that feeling *really* involves X, or it merely appears to you that it does¹⁵. Just going by the definition of zombies, a first-person analysis of X would also not enable you to determine which world you are from, let alone revealing what it is like to have an experience that merely seems as if it has X, and how that differs from an experience with X. It seems that both first-person and third-person analyses of X fall short in capturing its nature.

6. The intertwining of first-person and third-person

We can now ask once more, what is X? In essence, X, if real, remains elusive to both first-person and third-person inspections. It defies physical description, resists introspection, and evades mathematical articulation. The challenge is not merely our incapacity to discern the difference

¹⁵ The question is not whether you undergo an ineffable experience of pain. Rather, it is whether that ineffable experience involves X. I invite philosophers who believe it is possible for one to know one is not a zombie to explain how, without adhering to circular arguments or mere assertions.

between world one and world two; rather, the predicament lies in the inherent impossibility of distinguishing between them. Consequently, the two worlds (at least epistemically) collapse into one another, rendering the belief in inhabiting either one as practically inconsequential beyond personal satisfaction.

Let us revisit the initial question posed at the beginning of the essay: How does one come to learn something about consciousness? There is an aspect of consciousness that one can comprehend only through scientific experimentation or via philosophical reflection. Simultaneously, there is a facet of *my* consciousness, or of *your* consciousness, that is accessible only through personal exploration—sensed through practices such as meditation, psychedelic experiences, and the like. Much like someone congenitally deaf cannot grasp the experiential sensation of sounds from reading a book, one cannot know all there is to know about consciousness *semantically* or by understanding neuronal activity. This is a distinct issue from whether consciousness is physical or nonphysical and can therefore be considered another neutral point of agreement about conscious experience.¹⁶

A useful metaphor to highlight the first-person/third-person distinction is to think of conscious experience as an open-world game. One could learn a lot about the game simply by studying the code and yet, this looks nothing like playing the game itself. In fact, it is possible that certain game events are not hard-coded and that they would be missed if one were to only focus on the code. Conversely, one can relish in the pleasures of the gameplay, uncovering hidden secrets and uncharted territories, all while remaining oblivious to the underlying code. However, understanding coding techniques commonly used in game development can grant players unique insights into potential gameplay enhancements and strategic possibilities when tinkering with game mechanics. Ultimately, both playing the game and studying the code provide additional context for one another though they serve different purposes. Some gamers write code and some programmers play games, but gaming is not the same as programming. Where studying consciousness is concerned, it is crucial to refrain from interpreting claims of one domain as unequivocal truths about the other.¹⁷

7. Concluding remarks

The nature of consciousness remains both a topic of controversy and also a “meta-controversy”; philosophers find themselves at odds regarding the significance of certain areas deemed crucial by others. In this paper, I tried to show that adopting a more “meta” approach

¹⁶ This also differs somewhat from the “ability hypothesis” as one of the famous replies to the Mary thought experiment (Nida-Rümelin and O’Conaill, 2023). Unlike in Jackson’s thought experiment and its relevant discussions, my concern here is not whether one possesses all the physical facts about consciousness, or how a fine-grained notion of the term “knowledge” refers to subjective experience. Rather, it is simply to highlight that there are different domains encountered by first-person and third-person inspections, even if there is an overlap between them.

¹⁷ A discussion of games and code implies an underlying causal relationship that I do not intend to convey here. The analogy is merely regarding the distinct, and yet *correlational* relationship between the first-person and third-person perspectives. Whether or not the relationship is causal is another matter for discussion.

towards the nature of subjective experience and opting to find a neutral explanandum ultimately leads to an epistemic veil. More importantly, I argued that if science (and/or physicalism) truly falls short in grasping an important aspect of subjective experience, it is not clear how anything else can fill that gap either.

While I attempted to explore a neutral explanandum without specific metaphysical commitments, this endeavor itself can be perceived as a form of ontological commitment, resting on various assumptions. For example, granting the existence of zombies is not acceptable to some philosophers, and defending the possibility of illusionism is viewed as an oxymoron by others. I believe that the way out of this impasse is to examine different domains of first-person-based and third-person-based approaches on their own terms and within various ontologies, while avoiding, or at least *postponing*, the attempt to draw literal links between these domains. In practice, this implies treating different scientific (or otherwise) theories of consciousness less as theories about “consciousness-the-fact-of-the-world”, and more as explorations of how each particular view defines and treats its specific subject, which it refers to as consciousness.

This is similar to how an anthropologist would study different cultures’ treatment of God(s), without himself having to believe in any of them. In the context of Norse mythology, Odin *did* sacrifice an eye in his pursuit of knowledge and in the context of Catholic belief, Jesus *was* resurrected after crucifixion. The same goes for understanding theories of consciousness. In panpsychist frames of thought, consciousness simply *is* a fundamental part of the universe and in physicalist models, consciousness *is* the result of physical states and brain activity.

While it may seem that I am equating scientific models (that involve falsifiability and predictive power) to religious beliefs, this is not my intention. My claim is that if one adopts a meta position and merely *treats* theories of consciousness *as if* they are stories, this pretense and dissociation itself can provide us with useful information about consciousness as we go about contemplating different perspectives.

Furthermore, rather than rejecting formulations of consciousness that do not fall in line with one’s view, perhaps another approach could prove fruitful and that is to resist the temptation of a *hasty dismissal*, even for views that seem entirely crazy. One can simply make peace with the possibility (and its subsequent disappointment) that many of such frameworks will not adequately address what one would have wish them to, and instead, focus on what it is that these frameworks explore.

Needless to say, strong proponents of each view may perceive this approach as a waste of time since it is *their* framework that provides the best answers to questions *worth asking*. However, for others who harbor serious doubts regarding all these perspectives, this opens the door to a different strategy for reconciling the conflicting views. It is quite plausible that an ultimate theory of consciousness would involve integrating concepts from multiple paradigms that seem incompatible with one another at this moment.

8. Acknowledgment

I extend my sincere gratitude to Markus Eronen and Fred Keijzer for their comprehensive feedback on the entirety of this essay. Additionally, I wish to express my appreciation to Philip Goff and the late Daniel Dennett for their valuable remarks. Their constructive input has significantly enriched the quality of this work.

9. References

- Block, N. (2002). Some concepts of consciousness. *Philosophy of mind: Classical and contemporary readings*, 206-218.
- Chalmers, D. J. 1996a, Facing up to the problem of consciousness. *Toward a science of consciousness: the first Tucson discussions and debates*. MIT Press, Cambridge, 5-28.
- Chalmers, D. J. 1996b, *The Conscious Mind: In Search of a Fundamental Theory*, New York and Oxford: Oxford University Press.
- Chalmers, D. (2018). The meta-problem of consciousness.
- Dennett, D. C. (1988). Quining qualia. *Consciousness in contemporary science*, 42-77.
- Dennett, D. C. (1993). *Consciousness explained*. Penguin uk.
- Dennett, D. C. (1995). The unimagined preposterousness of zombies. *Journal of Consciousness Studies*, 2, 322-325.
- Dewhurst, J. (2021). Folk psychological and neurocognitive ontologies. *Neural mechanisms: New challenges in the philosophy of neuroscience*, 311-334.
- Feyerabend, P. (2020). *Against method: Outline of an anarchistic theory of knowledge*. Verso Books.
- Frankish, K. (2012). Quining diet qualia. *Consciousness and Cognition*, 21(2), 667-676.
- Frankish, K. (2016). Illusionism as a theory of consciousness. *Journal of Consciousness Studies*, 23(11-12), 11-39.
- Gade, M., Tan, L. L., Damry, A. M., Sandhu, M., Brock, J. S., Delaney, A., ... & Laurino, P. (2021). Substrate dynamics contribute to enzymatic specificity in human and bacterial methionine adenosyltransferases. *JACS Au*, 1(12), 2349-2360.
- Galilei, G. (1957). Il saggiaiore (The assayer). *The Controversy of the Comets of, 1618*.
- Goff, P. (2019). *Galileo's error: Foundations for a new science of consciousness*. Vintage.

Herzog, M. H., Doerig, A., & Sachse, C. (2023). Why computational complexity may set impenetrable barriers for epistemic reductionism. *Synthese*, 202(5), 136.

Jackson, Frank (1982). Epiphenomenal qualia. *Philosophical Quarterly* 32 (April):127-136.

Jacobellis v. Ohio, 378 U.S. 184 (1964) (Stewart, J., concurring).

Kammerer, F. (2021). The illusion of conscious experience. *Synthese*, 198(1), 845-866.

Kirk, Robert, "Zombies", *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/fall2023/entries/zombies/>>.

Lewis, David, 1983a, "Extrinsic Properties", *Philosophical Studies*, 44:197–200.

Marshall, Dan and Brian Weatherson, "Intrinsic vs. Extrinsic Properties", *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/fall2023/entries/intrinsic-extrinsic/>>.

Montero, B. (2001). Post-physicalism. *Journal of Consciousness Studies*, 8(2), 61-80.

Montero, B. G. (2012). Irreverent physicalism. *Philosophical Topics*, 91-102.

Nagel, T. (1974). What is it like to be a bat. *Readings in philosophy of psychology*, 83(4), 435. <https://doi.org/10.2307/2183914>

Nida-Rümelin, Martine and Donnchadh O Conaill, "Qualia: The Knowledge Argument", *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/fall2023/entries/qualia-knowledge/>>.

Peterson, R. T. (1947). *A Field Guide to the Birds*. Boston: Houghton Mifflin.

Piccolino, Marco, and Nicholas J. Wade, 'Galileo's sensory philosophy', *Galileo's Visions: Piercing the spheres of the heavens by eye and mind* (Oxford, 2013; online edn, Oxford Academic, 23 Jan. 2014), <https://doi.org/10.1093/acprof:oso/9780199554355.003.0010>

Rahimian, S. (2022). The myth of when and where: How false assumptions still haunt theories of consciousness. *Consciousness and Cognition*, 97, 103246.

Stoljar, Daniel, "Physicalism", *The Stanford Encyclopedia of Philosophy* (Spring 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), forthcoming URL = <<https://plato.stanford.edu/archives/spr2024/entries/physicalism/>>.

Sundström, P. (2014). Two types of qualia theory. *The Harvard Review of Philosophy*, 20, 107-131.

Tye, Michael, "Qualia", *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2021/entries/qualia/>>.