

Having Experience and Knowing Experience: A Case for Illusionism about Phenomenal Consciousness

Daniel Shabasson

Independent Scholar

Abstract

Most philosophers think that phenomenal consciousness is real and that it has two components: an experiential component—a state that is subjectively ‘like something’ for a subject of experience; and a cognitive component—the subject’s awareness of the experiential component and knowledge of what it’s like. Illusionists, by contrast, claim that phenomenal consciousness is an illusion. It does not exist but only seems to exist (Frankish 2016). Although illusionism is highly counterintuitive, I shall claim that it is probably true. For I shall argue that phenomenal realism—the view that phenomenal consciousness is real—is conceptually incoherent. I identify four possible realist pictures of phenomenal consciousness, individuated according to the stand they take on two fundamental questions about phenomenal consciousness: (i) whether the cognitive component would fallibly or infallibly represent the experiential component to the subject of experience; and (ii) whether a cogent picture of phenomenal consciousness must include both the experiential and the cognitive component, or whether it might be constituted by one component alone. I examine these four realist pictures of phenomenal consciousness and show that each of them is conceptually incoherent. Therefore, I argue, phenomenal realism is conceptually incoherent.

Keywords: Phenomenal Consciousness, Illusionism, Eliminativism, Qualia, Experience.

1. Introduction

Most philosophers think that phenomenal consciousness is real and that it has two components: an experiential component—a state that is subjectively ‘like something’ for a subject of experience (Nagel 1974); and a cognitive component—the subject’s awareness of the experiential component and knowledge of what it’s

like.¹ Illusionists, by contrast with phenomenal realists, claim that phenomenal consciousness is an illusion. It does not exist but only seems to exist (Frankish 2016). Illusionism is highly counterintuitive and has been called the silliest claim ever made (Strawson 2018). Nevertheless, I shall claim that illusionism is probably true. For I shall argue that phenomenal realism—the view that phenomenal consciousness is real—is conceptually incoherent.

My argument proceeds as follows. In §2, I identify four possible realist pictures of phenomenal consciousness, individuated according to the stand they take on two fundamental questions about phenomenal consciousness: (i) whether the cognitive component would *fallibly* or *infallibly* represent the experiential component to the subject of experience; and (ii) whether a cogent picture of phenomenal consciousness must include both the experiential and the cognitive component, or whether it might be constituted by one component only—either the cognitive component alone or the experiential component alone. In §3-§6, I examine these four realist pictures of phenomenal consciousness and argue that each of them is conceptually incoherent. Therefore, I argue, phenomenal realism is conceptually incoherent. In §7, I argue that illusionism avoids the worries I raise for phenomenal realism, and I compare the relative merits of illusionism and mysterianism (McGinn 1989) in addressing those worries.

2. Four Realist Pictures of Phenomenal Consciousness

Phenomenal realists maintain that phenomenal consciousness is real and generally believe it has two components. The first is its experiential component. For example, it is supposed to be subjectively *like something* (Nagel 1974) to visually experience the redness of a ripe strawberry. Philosophers have used various terms for this experiential component: a ‘quale’, a ‘phenomenal property’, a ‘what-it’s-likeness’, a ‘phenomenally conscious experience’, a ‘subjective experience’, or just plain ‘experience’. It is a mental state (or property) subjectively like something for the subject who instantiates it. I shall call this an ‘experiential state’.

Experiential state A mental state (or property) that is subjectively *like something* for the subject instantiating it.

The second component of phenomenal consciousness is its cognitive component—the subject’s awareness of her experiential state and her knowledge of what it’s like. For example, consider a subject visually experiencing the redness of a ripe strawberry, who would instantiate a reddish experiential state. She would be aware of this experiential state and know that it has a reddish (as opposed to a greenish or bluish) phenomenal character. And she would know what it’s like subjectively to instantiate an experiential state with a reddish phenomenal

¹ As Joseph Levine expressed the relation between the experiential and cognitive components: “Qualia are such as to necessitate awareness of them” (Levine 2001: 168). As Phillip Goff expressed the relation: “A pain is a feeling, and all there is to a feeling is how it feels. And when you feel it, you know how it feels” (Goff 2020: at 1:05:58). As Chalmers expressed the relation, “there is something intrinsically epistemic about experience. To have an experience is automatically to stand in some sort of intimate epistemic relation to the experience. [...] There is not even a conceptual possibility that a subject could have a red experience like this one without having *any* epistemic contact with it: to have the experience is to be related to it in this way” (Chalmers 1997: 196-197).

character. Call this cognitive component of phenomenal consciousness a ‘cognitive state’.

Cognitive state The subject’s awareness of her experiential state and knowledge of what it’s like.²

Following Russell, many phenomenal realists—proponents of the ‘direct acquaintance view’—maintain that our knowledge of our experiential states (i.e., our cognitive states) is a form of *knowledge by acquaintance* (Russell 1910-11). Some proponents of the direct acquaintance view claim that this knowledge is *non-propositional*, i.e., a direct and immediate relation between an agent and an object or a property (Duncan 2023; Giustina 2021). Presumably, the non-propositional nature of this knowledge would account for its *ineffability*—the impossibility of explaining what it’s like to see red to a person who has never seen red (Jackson 1982). Other proponents of the view maintain that acquaintance involves the subject’s deployment of “direct phenomenal concepts” yielding *propositional* knowledge of experiential states (Chalmers 2003: 235). On either view, direct acquaintance is supposed to be a non-causal relation, meaning that the subject of experience would automatically and infallibly know the nature of her experiential state merely by virtue of its instantiation in her. This contrasts with causal views of knowledge according to which we acquire knowledge through fallible causal processes such as representation. Although I do not ascribe to either variety of the direct acquaintance view or accept the existence of non-propositional knowledge, for the sake of argument I shall remain fully agnostic about the nature of the cognitive state. I take no stand on whether the object of the subject’s knowledge would be a proposition, an object, or a property; or whether the subject’s access to this knowledge would be grounded in direct acquaintance or mediated by causal representational processes in the brain.

The direct acquaintance view typically goes hand in hand with the thesis that it would be metaphysically impossible for a cognitive state to present an experiential state non-*veridically*. Our experiential states are said to *reveal* themselves to us—i.e., their natures are transparent to us in such a way that we cannot fail to see them as they truly are. As Philip Goff summed up the thesis of revelation, “A pain is a feeling, and all there is to a feeling is how it feels. And when you feel it, you know how it feels” (Goff 2020: at 1:05:58). (See Chalmers 2018: 25 for a discussion of revelation.) Other realists about phenomenal consciousness, some of whom may not overtly ascribe to the direct acquaintance view, are committed to the similar ‘no-gap’ thesis, which posits that when it comes to phenomenal consciousness, one cannot sensibly draw a distinction between appearance and reality (Kripke 1980). As Strawson expressed the idea, “to seem to feel pain is to

² The relationship between the experiential state and the cognitive state will be an important theme in this paper. I often speak of these states as if they were separate and distinct components of phenomenal consciousness. At the same time, these components are *intertwined* with one another in an idiosyncratic way: the instantiation of a cognitive state seems to be a necessary condition for the instantiation of an experiential state. For it seems impossible to conceive of a state that is subjectively like something for a subject without the subject’s awareness of it. So I sometimes speak of a cognitive state as if it were an *essential property* of an experiential state, which suggests that the cognitive state is a component of the experiential state, rather than a component of phenomenal consciousness. These confusions are unavoidable, as one of the central arguments of the paper, especially in §3 and §6, is that there is no way to fit the cognitive and experiential states together coherently.

be in pain. It's not possible here to open up a gap between appearance and reality, between what is and what seems" (Strawson 2018: 1). (See also Searle 1998: 1941; Moran 2001: 14; Horgan 2012: 406.) Like proponents of the direct acquaintance view, no-gap theorists see the possibility of any discrepancy between the nature of the experiential state and the way that state is presented to the subject of experience to be inconceivable.

The intuition motivating the direct acquaintance view and the no-gap thesis is fundamentally the same. If it seems to me that I am experiencing a red afterimage, it is deeply difficult to make sense of the claim that, unbeknownst to me, I might be experiencing a green afterimage. Intuitively, I am authoritative about my subjective experience: if an afterimage appears to me to be red, then it simply *is* red. Call this the 'infallibility intuition'.

The infallibility intuition The intuition that we are authoritative about our experiential states and know infallibly what they are like.

I understand the appeal of this intuition, but I do not trust it. The history of philosophical and scientific inquiry has revealed that powerful intuitions can be erroneous. Furthermore, there is compelling empirical evidence that our first-person introspective judgments about our occurrent mental states are sometimes mistaken in surprising ways (Dennett 1993, 2006). At the same time, I shall not presuppose that the infallibility intuition is erroneous. For the sake of argument, I shall take seriously both the view that non-veridical cognitive states are metaphysically possible and the view that they are not.

I shall use 'fallible' and 'infallible' to describe contrasting views about the reliability of cognitive states, as follows:

Fallible cognitive states Cognitive states are *fallible* iff it is metaphysically possible that they are non-veridical. There is at least one possible world in which it is nomically possible for a cognitive state to present an experiential state non-veridically.³

Infallible cognitive states Cognitive states are *infallible* iff it is metaphysically impossible that they are non-veridical. There is no possible world in which it is nomically possible for a cognitive state to present an experiential state non-veridically.

By 'fallibility world', I mean the following:

Fallibility world A fallibility world is a possible world in which it is nomically possible for a cognitive state to present an experiential state non-veridically.

By 'infallibility world', I mean the following:

³ For expository convenience, I have defined cognitive states as *knowledge* states. But this needs to be qualified. Instantiating a non-veridical cognitive state would not constitute knowledge, for knowledge is a *factive* state. The reader should suppose that non-veridical cognitive states, if metaphysically possible, would be *non-factive* attitudinal states. If it turned out that the content of cognitive states was *propositional*, a non-veridical cognitive state would be a *false belief*. If, on the other hand, the content of cognitive states turned out to be *non-propositional*, a non-veridical cognitive state would be a token of *the non-propositional analog of false belief* (for which there is no standard term of art), since belief is generally strictly defined as an attitudinal relation between an agent and a proposition, but not between an agent and any sort of non-propositional entity.

Infallibility world An infallibility world is a possible world in which it is nomically impossible for a cognitive state to present an experiential state non-*veridically*.

Proponents of the view that cognitive states are infallible (most direct acquaintance and no-gap phenomenal realists) would deny the existence of fallibility worlds. All possible worlds would be infallibility worlds. Proponents of the view that cognitive states are fallible (some phenomenal realists, especially those committed to physicalism) would accept the existence of fallibility worlds.

Thus far, I have presupposed that any coherent picture of phenomenal consciousness would have *two* components. There's what it's like (an experiential state), and then there's knowing what it's like (a cognitive state). Of course, it's an open possibility that a philosopher might propose that a cogent picture of phenomenal consciousness would include *one* component only—either a cognitive state alone or an experiential state alone. It is not clear whether any philosopher has actually proposed either sort of one-component picture, but I shall be thorough and consider all possible views.

One-component picture of phenomenal consciousness

Phenomenal consciousness would be constituted by one component only—either a cognitive state alone or an experiential state alone.

Two-component picture of phenomenal consciousness

Phenomenal consciousness would be constituted by both experiential and cognitive states.

In light of the two broad factors relevant to taxonomizing the possible pictures of phenomenal consciousness I have discussed in this section—(i) whether cognitive states are fallible or infallible, and (ii) whether a cogent picture of phenomenal consciousness might include one component alone or must include both components—there are four possible pictures of phenomenal consciousness, as follows:

- (1) **Two-component picture:** experiential states with *fallible* cognitive states (discussed in §3)
- (2) **One-component picture:** cognitive states alone (discussed in §4)
- (3) **One-component picture:** experiential states alone (discussed in §5)
- (4) **Two-component picture:** experiential states with *infallible* cognitive states (discussed in §6)

Every theory of phenomenal consciousness will need to take a stand on the two questions that serve as the individuation criteria for these four pictures. I believe, therefore, that these four pictures are broad enough and general enough for any theory of phenomenal consciousness, regardless of its specifics, to fit squarely into the ambit of one of them. In §3-§6, I shall discuss these four pictures in the order listed above and argue that each is conceptually incoherent. Therefore, I argue, phenomenal realism is conceptually incoherent.

3. Two-component Picture: Experiential States with *Fallible* Cognitive States

Suppose that cognitive states were fallible, i.e., it was metaphysically possible for them to present experiential states to a subject of experience non-*veridically*. This is the sort of picture that no-gap and direct acquaintance realists would reject, for they take cognitive states to be infallible. I shall claim that no-gap and direct

acquaintance realists are *right* to oppose this picture. I argue—via a thought experiment I call the *pain-pleasure inversion chamber* (or ‘PPIC’)—that this picture of phenomenal consciousness faces a worry, which I call ‘the otiosity problem’, that shows that the picture is conceptually incoherent.

The PPIC is a high-tech device developed by intelligent beings in a hypothetical fallibility world. It has space for one occupant, with numerous wires attached to her brain. The occupant is in the chamber for ten minutes during which time the machinery in the chamber manipulates her brain to cause her cognitive states to present her experiential states to her non-veridically. The device has two settings. On the *cognitive-pleasure* setting, she is continuously in an excruciatingly painful experiential state but instantiates a non-veridical cognitive state that continuously presents her painful experiential state to her as highly pleasurable. After exiting the chamber, she recalls the experience as pleasurable and reports it as such. On the *cognitive-pain* setting, the inverse takes place. For ten minutes, the occupant is continuously in an intensely pleasurable experiential state but continuously instantiates a non-veridical cognitive state that presents her experiential state to her as excruciatingly painful. After exiting the chamber, she recalls the experience as excruciatingly painful and reports it as such. On either setting, the subject’s cognitive state is 180 degrees mistaken about the nature of her experiential state.

Pain-pleasure inversion chamber (‘PPIC’) thought experiment

The PPIC device causes the subject’s cognitive states to be continuously 180 degrees mistaken about the nature of her experiential states.

Now suppose that you have decided to try out the PPIC chamber for yourself. Here is the pivotal question: *Which setting would you choose to ensure a pleasurable time and avoid a painful one—the cognitive-pleasure or the cognitive-pain setting?* What would determine what your experience in the chamber would be like for you—your experiential state or your cognitive state that presents it non-veridically?

I’d be willing to bet you would select the cognitive-pleasure setting on the PPIC device, electing to be in an experiential state of excruciating pain that your cognitive state presents as pleasurable. If you are skeptical about this, consider which setting you would select if you had *previously* tried out the PPIC chamber on the cognitive-pain setting. You would recall your experience in the chamber as excruciatingly painful. In selecting the setting on the device, you would care only about how your cognitive state presents your experiential state and not one whit about the nature of the experiential state. While in the chamber on the cognitive-pleasure setting, you would be glad you had chosen that setting. Hence, you would answer the pivotal question above in favor of the cognitive state alone mattering to you. What it’s like for you would be a function exclusively of the nature of your cognitive state. The experiential state would be *otiose*, doing no philosophical work, and it ought to drop out of this two-component realist picture of phenomenal consciousness. Call this the ‘otiosity problem’.

The Otiosity Problem If your cognitive states were fallible, experiential states would be *otiose*, explanatorily idle posits. What it’s like for you would be a function exclusively of your cognitive states. Doing no philosophical work, experiential states ought to drop out of the picture.

Instantiating a painful experiential state would not be necessary for being in pain—for you could be in pain when you instantiate an experiential state of

pleasure, but your cognitive state presents it as painful (e.g., on the cognitive-pain setting). Nor would instantiating a painful experiential state be sufficient for being in pain—for you could instantiate a painful experiential state but feel pleasure because your cognitive state presents that painful experiential state as pleasurable (e.g., on the cognitive-pleasure setting). Since instantiating a painful experiential state would be neither necessary nor sufficient for being in pain, instantiating a painful experiential state \neq being in pain. Furthermore, instantiating a cognitive state presenting your experiential state as painful would be both necessary and sufficient for being in pain. So the instantiation of a cognitive state presenting your experiential state as painful would = being in pain. Likewise, instantiating a cognitive state presenting your experiential state as pleasure would = experiencing pleasure. What it's like for you would be a straight function of the nature of your cognitive state. The what-it's-likeness would be located in the cognitive state, rather than in the experiential state. The experiential state, which was supposed to *be* the what-it's-likeness, would no longer play any role in the picture.

To be clear, the otiosity problem would arise without regard to whether a subject happened to inhabit a fallibility world or an infallibility world. The PPIC thought experiment shows that your cognitive state presenting things to you as painful is necessary and sufficient for being in pain, so being in a cognitive state presenting pain to you *is* to be in pain. It's just that in some worlds—in infallibility worlds, a painful cognitive state would be regularly accompanied by a painful experiential state as a matter of nomic—but importantly, not metaphysical—necessity. Your painful cognitive state's being accompanied by a painful experiential state would be an *accidental property* of pain. In deciding which setting to select on the PPIC device to avoid pain, you would waste your time wondering whether your cognitive states presented your experiential states veridically or non-veridically, or whether you inhabited a fallibility or an infallibility world. In every possible world, whether in a fallibility or an infallibility world, you would guarantee experiencing pleasure and avoiding pain by choosing the cognitive-pleasure setting for the PPIC device. The experiential state would have been stripped of its role in the picture—to be the bearer of the 'what-it's-likeness', which would now be borne solely by the cognitive state.

The experiential state would do no philosophical work and ought to be removed from the picture,⁴ leaving us with a one-component picture consisting

⁴ An anonymous reviewer objected that despite their otiosity, we might retain experiential states in this picture and consider them to be epiphenomenal—causally inefficacious, but nevertheless real. However, the question of whether experiential states are epiphenomenal is orthogonal to the issue of their otiosity. The otiosity issue is whether experiential states play any explanatory role in the picture. Many theories of phenomenal consciousness hold that phenomenal properties are epiphenomenal, yet maintain that they play important explanatory roles. For example, direct acquaintance realists often suppose that phenomenal properties are epiphenomenal, yet they play a central role in the theory. The phenomenal properties are the what-it's-likenesses, which despite being epiphenomenal, i.e., causally inefficacious, exist and are known to the subject of experience through acquaintance. I have argued that experiential states would, on this two-component picture of phenomenal consciousness, play no explanatory role, i.e., they would otiose. The experiential states were meant to play the role of the what-it's-likenesses, but the PPIC shows that the what-it's-likenesses would reside in the cognitive states instead. Stripped of their intended theoretical purpose—to be the states that are like something for the subject to instantiate them—they lack any explanatory role.

solely of the cognitive state (which I discuss in §4 below). Therefore, I conclude, this two-component picture fails to characterize phenomenal consciousness. Again, realist proponents of the direct acquaintance view and no-gap thesis should agree with this conclusion, given that this two-component picture is predicated on the existence of non-veridical cognitive states, the metaphysical possibility of which direct acquaintance and no-gap phenomenal realists would deny.

4. One-Component Picture: Cognitive States alone

As I have argued above, removing experiential states from our two-component picture to cure the otiosity problem would leave us with a one-component picture with the cognitive state as its sole component. But such a one-component picture of phenomenal consciousness would not be coherent either. We would now confront a worry about the cognitive state, which I call ‘the intentionality problem’. The content of a cognitive state is supposed to be knowledge. Knowledge is always *about* something, whether the object of knowledge is a proposition or a non-propositional entity such as an object or property, or whether the knowledge is accessed via representational mechanisms in the brain or through direct acquaintance. A cognitive state without intentionality—one that was not about anything—would lack any truth-evaluable content. The content of a cognitive state presenting pain would be schematized as: _____ *is painful*, which would lack truth-value due to the empty argument place. The cognitive state would not attribute pain *to* any experiential state or *to* anything or *to* anyone. A subject instantiating such a cognitive state would not believe that she, or anyone in particular, was experiencing pain.

The Intentionality Problem If the experiential state has dropped out of the picture, what would the cognitive state be about? To what entity or individual would the cognitive state attribute pain? If the cognitive state lacked an intentional object, it would be about nothing and lack any truth-evaluable content.

Perhaps, a phenomenal realist might propose, we might avoid the intentionality problem if the painful cognitive state was *about itself* (instead of a painful experiential state). The cognitive state would be *self-referential*, serving as its own intentional object. It would say of itself that it is painful for the subject of experience to instantiate that very cognitive state. Its content might be schematized (from the first-person perspective) as follows:

Content of a self-referential cognitive state

The instantiation of this very cognitive state, which presents pain, is painful for me.

However, this proposal, contrary to initial appearances, does not work. The above schematization of content is *sub silentio* a cognitive state about an experiential state, despite the misleading absence of the expression ‘experiential state’. Recall that I defined an experiential state as a state that’s like something to instantiate it. The above schematization of content says that the cognitive state of pain I am currently instantiating is like something for me to instantiate it—that it is painful for me to instantiate it. So the schematization says that for me to instantiate a painful cognitive state *is* for me to instantiate a painful experiential state. We could, therefore, more transparently reword the schematization in the following way to reveal that it tacitly refers to an experiential state:

Revised content of a self-referential cognitive state

The instantiation of this very cognitive state, which presents pain, *is a painful experiential state for me.*

The third clause of the original schematization is now worded in the revised schematization in terms of a painful experiential state (via the language in italics), making it evident that the original wording ('is painful for me') refers to a state that is painful for me to instantiate, which, by the definition of 'experiential state', is the same thing as me instantiating a painful experiential state.

We now confront the claim, in the italicized third clause of our revised schematization, that a painful cognitive state 'is' a painful experiential state. There are two possible ways to unpack the meaning of this 'is'. On a first interpretation, being in a painful cognitive state would be *strictly identical* to being in a painful experiential state. But this interpretation is a non-starter. An experiential state is supposed to be a subjective phenomenal feel, a what-it's-likeness. A cognitive state is the subject's awareness of and knowledge of that what-it's-likeness. We can make no sense of the claim that these are identical any more than we could make sense of the claim that the proposition that snow is white is identical to knowing that snow is white.⁵ On a second interpretation of 'is', the revised schematization would say that being in a painful cognitive state *gives rise to* a painful experiential state for the subject. The cognitive and experiential states would be distinct entities—the former one generating the latter. This is a far more plausible interpretation of the 'is', but problematically, it brings us back to a *two-component* picture consisting of both experiential states and cognitive states. Such a two-component picture is precisely what we wanted to avoid, for, assuming the cognitive state was fallible, such a picture would confront the otiosity problem (as I argued in §3 above). And if the two-component picture were one in which the cognitive state was *infallible* about the experiential state, it would be the sort of two-component picture I examine below in §6 and argue is incoherent there. My point is that positing a self-referential cognitive state to avoid the intentionality problem—to provide the cognitive state with an intentional object—would end up delivering a two-component picture, thus failing to save this one-component picture constituted by cognitive states alone. Unless there is some other way to avoid the intentionality problem, this picture of phenomenal consciousness is incoherent.

5. One-Component Picture: Experiential States alone

I shall argue that a one-component picture constituted solely by experiential states would not be a coherent picture of phenomenal consciousness. Experiential states lacking cognitive states about them would not be like anything for the subject instantiating them. That is, the subject's awareness of and knowledge about her experiential states would be an essential property of them, without which they

⁵ My point is that knowing a proposition *p* or knowing some object *o* is always going to be distinct from the entity known, *p* or *o*. And this point can also be made just as well in terms of sentences, rather than propositions. Any utterance (in English) of the form 'A is Φ ' will always mean something different than any utterance of a sentence of the form 'I know that A is Φ '. And the same point should apply to cognitive and experiential states whether we consider cognitive states and experiential states as state *types*, or we consider them as state *tokens*, as dated particulars.

would not count as experiential states. As Joseph Levine expressed this idea (referring to what I call “experiential states” as “qualia”): “Qualia are such as to necessitate awareness of them” (Levine 2001: 168). Levine finds this feature of qualia—that they are invariably objects of awareness—paradoxical. He writes:

Awareness certainly seems to be a relation, which would entail that one can distinguish the act from the act from the object of awareness. Yet when it comes to qualia, to the contents of conscious experience, the two don’t come apart so easily. It does seem impossible to really separate the reddishness from the awareness, yet it also seems impossible to tell a coherent story about how this could be so. I wish I had the right story to tell; my aim is to press the depth and urgency of the need for such a story (Levine 2001: 9).

We cannot, in Levine’s words, “separate” a quale from our awareness of it because the awareness would be an integral and essential component of the quale itself, without which the quale would not exist. No awareness, no quale.

Frankish has recently articulated a related (although somewhat broader) point about the connection between experiential states and our awareness of them:

We make some assumptions about these states [conscious states]. We assume that we are *aware of them* (it’s implicit in calling them ‘conscious’). We assume that their natures matter to us (some would say it’s the only thing that matters to us). And we assume that they have a causal influence on our behavior (what we experience affects what we do). These assumptions are plausibly connected; our conscious experiences influence our behavior because they matter to us, *and they matter to us because we are aware of them* (Frankish 2021b: 65; all italics mine).

According to Frankish, when it comes to phenomenal consciousness, we’re talking about states that we are conscious of, i.e., states we are aware of. After all, we’re talking about phenomenal *consciousness*, not phenomenal *unconsciousness*. Moreover, if experiential states didn’t affect us psychologically or behaviorally, and we were not aware of their existence, we could not *care* about them. But, according to phenomenal realism, we care about them very much. Hence, we must be aware of them. Putative experiential states not accompanied by cognitive states would not be ‘like anything’ for the subject instantiating them, so not experiential states after all.

To further explore examine this question—whether cognitive states would be necessary for experiential states, consider a thought experiment proposed by Keith Frankish featuring a being he calls a “representational zombie” (which I shall refer to as an ‘r-zombie’ for short), a being with experiential states but no cognitive states about them. According to Frankish, an r-zombie would lack

cognitive access to its phenomenal properties [which here, I call ‘experiential states’] and would be unable to form beliefs about them, reflect on them, report them, remember them, respond emotionally to them, or act upon them (Frankish 2016: 13).

Lacking cognitive states, an r-zombie would be unaware that she instantiated any experiential states. Nor would she know anything about their distinctive what-it’s-likenesses—the distinctive subjective raw feel or phenomenal character

of each type of experiential state differentiating it from other types. If you asked an r-zombie what it's like to see red and how it's different from what it's like to see green, she would have no idea what you're talking about. In case it is not already obvious why the r-zombie would count as a sort of zombie, I'll discuss four reasons militating in favor of the view that the r-zombie would be a zombie and lack phenomenal consciousness. By showing that the r-zombie is a kind of zombie, I aim to demonstrate that what the r-zombie lacks—cognitive states—would be necessary for phenomenal consciousness (if it existed). And therefore, this one-component picture lacking cognitive states is incoherent.

First, the r-zombie would be completely indifferent towards her experiential states. We would usually say that our phenomenal consciousness matters to us. In fact, many phenomenal realists say that being phenomenally conscious is what makes life worth living. (Frankish: *ibid.*) However, a r-zombie would be incapable of caring about her experiential states. You cannot care about what you are unaware of, do not know about, or cannot find out about. Consider a short thought experiment to further pump this intuition, 'Old versus new anesthetic'.⁶ You are planning to undergo a painful knee surgery and you can choose between two types of local anesthetic. The old anesthetic prevents you from feeling pain by stopping painful experiential states from coming into being. By contrast, the new anesthetic allows experiential states to come into being, but it prevents you from becoming aware of them. It blocks your cognitive states from presenting your experiential states to you as painful (or as presenting any phenomenal character to you whatsoever). Clearly, it will make no difference to you whether you receive the first or the second anesthetic. You could not tell the anesthetics apart from the first-person perspective, so you could not care which one you received. This short thought experiment bolsters the claim that Frankish's r-zombie would not and could not care about her experiential states. From the first-person perspective, instantiating an experiential state of pain that you are not aware of and know nothing about is no worse than or distinguishable from having no experiential state of pain in the first place.

Second, without awareness of and knowledge about the experiential states she instantiates, a r-zombie would not grasp the concept of *experiential state* or *phenomenal consciousness*. According to the standard phenomenal realist intuition pumped in the knowledge argument (Jackson 1982), Mary the color scientist would need to instantiate a reddish experiential state to discover its distinctive phenomenal character—what it's like to see red as opposed to seeing green or feeling pain. She cannot find this out by reading a book on the nature of human color vision, no matter how sophisticated and comprehensive the book might be. But it is tacitly supposed in the thought experiment that it would be insufficient for Mary to learn what it's like to see red merely by instantiating a reddish experiential state without her noticing its presence. After all, it's hard to see how one could learn anything about an object in one's environment if one did not notice it. For example, the mere presence of a cat in my proximity would be insufficient for me to know, via first-person apprehension, that it is present or to learn anything about it. I would, at a minimum, have to be aware of the cat and attend to it. Likewise, Mary would have to be aware of her reddish experiential state and attend to it to learn what it's like to see red. And the same goes for what it's like

⁶ This is a modified version of a similar thought experiment by Keith Frankish (2021a: 1:15:15).

to see green, what it's like to feel pain, or what it's like to hear the sound of middle C. Without awareness of *any* of her experiential states, the r-zombie would not understand what it's like subjectively to see or feel anything whatsoever. She would not grasp the concept of *experiential state* or *phenomenal consciousness*, think that phenomenal consciousness existed, believe that there was any hard problem of consciousness, or have any problem intuitions about phenomenal consciousness (such as, e.g., that was it was non-physical or ineffable).

Consistent with the intuition pumped in the knowledge argument, phenomenal realists often suppose that there is a close relationship between being phenomenally conscious and understanding what phenomenal consciousness is. We acquire the concept of phenomenal consciousness only through our direct experience of our own experiential states, and not through mere conceptual understanding. So Mary learns what it's like to see red only by directly experiencing it from the first-person perspective. And conversely, there is the intuition that you will know what it means to be phenomenally conscious once you have seen for yourself what it's like to be phenomenally conscious from the first-person perspective. When Mary finally sees red, she comes to know what it's like to see red. If she didn't learn what it's like to see red when she saw a red flower for the first time, we'd have to conclude that she was a zombie. In this vein, Schneider (2019) has argued that we might develop a modified Turing test for phenomenal consciousness in an AI that would take an AI's puzzlement about phenomenal consciousness—it's possession of certain "problem intuitions about phenomenal consciousness (Chalmers 2018), e.g., that it generated a hard problem, that it was non-physical, ineffable, or private—as evidence that the AI was phenomenally conscious. The idea behind her proposal, consistent with the intuition pumped in the knowledge argument, is that an AI could be puzzled by phenomenal consciousness only if it were phenomenally conscious. It could not grasp the concept of phenomenal consciousness, no matter how conceptually sophisticated and intelligent it was, without first-person familiarity with its own phenomenally conscious states. The r-zombie would fail Schneider's test for phenomenal consciousness, suggesting that she is a zombie. The r-zombie's lack of awareness of her experiential states means that she would not realize they were instantiated in her, nor could she learn what they were like. She could not acquire the concept of phenomenal consciousness from them nor possess Chalmers' problem intuitions about phenomenal consciousness.⁷ The upshot is that the r-zombie's lack of awareness of her experiential states means she would not know what it was like to experience *anything whatsoever*, and therefore the concept of phenomenal consciousness would be alien to her. And this strongly suggests that she would be a kind of zombie due to her lack of cognitive states.

Third, an r-zombie and a phenomenally conscious human would differ radically in their sensitivities and behavioral reactions to their experiential states. Without the r-zombie's awareness of a painful experiential state, it would have no impact on her psychologically. She would not report pain or exhibit any pain

⁷ Schneider's test is meant to be a *sufficiency* test for consciousness in an AI. The fact that an AI was puzzled about consciousness would be sufficient to establish its possession of phenomenal consciousness. But it would not be a necessity test, for we should expect some phenomenally conscious AIs to fail the test. After all, some creatures, such as dogs or young children, might be phenomenally conscious but lack the intellectual sophistication or language skills to pass the test.

behavior. Of course, she might report pain and behave as if she were in pain by coincidence. However, her pain behavior and verbal report of pain would be entirely explicable by the painful *functional psychological states* she instantiated, which would happen by chance to reflect the painful character of her painful experiential state. Without awareness of your painful experiential state, it would make no difference to what you would say or do. You would neither report pain, nor act like you were in pain, nor believe you were in pain. That's simply not being in pain.

Fourth, there's the question of ownership—to *whom* would the r-zombie's experiential states *belong*? With phenomenally conscious beings such as us, phenomenal realists might suppose that experiential states would belong to whichever subject was connected to them through awareness and knowledge of them. But r-zombies would lack this connection to them. Without being affected by these experiential states, and without awareness or knowledge of them, the question arises: on what principled basis could we say that an experiential state *belonged* to one r-zombie in particular rather than any other (or to no one at all)? It is doubtful that the experiential state being instantiated *inside her skull* would be sufficient to make it hers. The r-zombie would be no more perturbed by its instantiation inside her own skull than by its instantiation in someone else's. The r-zombie's experiential states would be, as Frankish has expressed it, 'free-floating' experiential states (2021: at 1:22:56), belonging to no one in particular. But it is absurd to suppose that there might be a subjective experience with no subject in particular experiencing it.⁸

In summary, I contend that a subject's awareness of and knowledge about her experiential states would be the glue required to connect her experiential states to her world and make them hers. Without cognitive states, a subject would not know about her experiential states or care about them. She would not understand what phenomenal consciousness was. She would not know what it was like to see red or feel pain. Cognitive states, through which a subject would be aware of her experiential states and would know what they are like, would be an essential component of phenomenal consciousness (if it existed). So this one-component picture lacking cognitive states fails.

6. Two-component Picture: Experiential States with *Infallible* Cognitive States

The fourth and final picture of phenomenal consciousness I shall consider is a two-component picture—experiential states with *infallible* cognitive states. An experiential state of pain would, in every possible world, be accompanied by a cognitive state veridically presenting that experiential state of pain as such. *Prima facie*, such a picture looks like it could avoid the problems faced by the three other pictures I discussed in §3-§5. On this fourth picture—in contrast with the two-

⁸ Since *no* r-zombie has cognitive states, *no* r-zombie could be connected to any experiential state first-personally. Experiential states could be detected and understood by r-zombies, if at all, only through third-person science (if it makes sense to say that experiential states would be amendable to scientific study at all). An r-zombie scientist in possession of third-person conceptual understanding of experiential states would not be connected to them in such a way such that we might say that *she* was the experiencer and the experiential state *hers*. That r-zombie scientist would view the experiential state only from the third-person perspective. I take it that experiences are supposed to be intrinsically first-personal with respect to their experiencer or 'owner'.

component picture presented in §3—it would be metaphysically impossible for a cognitive state to be non-veridical, so we could not coherently formulate the PPIC thought experiment described in §3. In avoiding the PPIC, we avoid the otiosity problem for the experiential state. We would also forestall the intentionality problem for the cognitive state, discussed in §4, which would arise only as a consequence of the otiosity problem. Moreover, unlike the one-component pictures discussed in §4 and §5, each of which we found wanting for omitting one essential component of phenomenal consciousness, this fourth picture would include both components—experiential and cognitive states. In short, we would have both cognitive and experiential states in the picture, and we would be free and clear of the otiosity and intentionality problems. So this picture should seem promising to phenomenal realists. It accords with the direct acquaintance and no-gap views, which posit that non-veridical cognitive states are metaphysically impossible.

However, this picture is incoherent, just as the other three pictures discussed in §3-§5. I argue below in §6.1 and §6.2 that, according to the constraints of this fourth picture, it would be logically impossible to define an experiential state except in a viciously circular manner. So this picture is incoherent.

6.1. The Two Essential Properties of an Experiential State, according to this Fourth Picture

On this fourth picture, a subject necessarily would know of any experiential state she instantiated that it was one. Here's why. On this picture, every experiential state would be necessarily accompanied by a veridical cognitive state through which the subject would be aware of that experiential state and infallibly know what it was like. Now, knowing what your experiential state is like entails that you know that it is like *something*. And an experiential state just is, by definition, a state that is like something. So a subject who knew that her experiential state was like something would necessarily know that it was an experiential state.⁹ On this fourth picture, it would be an *essential property* of every experiential state that its subject would know it was one, for according to this picture it would be metaphysically impossible for an experiential state to exist in the absence of a cognitive

⁹ An anonymous reviewer objected that a subject would not necessarily realize that her experiential state was one just by instantiating one. That would require an additional step: the subject would have to *categorize* her experiential state as one according to some theoretical framework. After all, the expression 'experiential state' looks like a theoretical term, a philosophical term of art. It is implausible to suppose that any subject of experience would necessarily realize that her experiential state fit under that theoretical classification. She would merely realize that the state was like something for her, but she would not say to herself, 'this is an experiential state'. However, the objection misses the mark. Property K can be stated purely in terms of what-it's-likenesses, as follows:

K: *The subject's knowledge of any state that is like something for her that it is like something for her.* This re-formulation of K in terms of 'what it's like' does not entail that the subject even knows what 'what it's like' means or that she speaks a language and can say 'it's like something for me'. No theoretical classification or speech act is required for the subject to know this, for according to this fourth picture, a subject of experience knows what her experiential state is like just by instantiating it, and this entails she knows it's like something. It's true that when property K is phrased in terms of 'experiential states', it can have the misleading feel of a technical notion. I used 'experiential state' for ease of exposition, but the argument could have been made just as well using the 'what it's like' locution.

state veridically presenting it to the subject of experience. Call this essential property of experiential states ‘K’ (for ‘knowledge’).

Essential property K The subject’s knowledge of any experiential state she instantiated that it was an experiential state.

It is plausible to suppose that an experiential state would bear *other* essential properties apart from K, even if we might not know what they might be at present. We can call the complete set of all these other essential properties, those in addition to essential property K, ‘R’ (for the ‘rest’ of the properties).

Essential property R The complete set of other essential properties of an experiential state, besides K, whatever they might turn out to be.

Given our rudimentary understanding of phenomenal consciousness (supposing it existed), we would likely be unable to identify most, if not all, of the properties included in R. However, it will not matter to my argument which properties would be included in R or whether they were discoverable. For whatever they might turn out to be, I will show in §6.2 that the inclusion of property K among the essential properties of experiential states would, on its own, be sufficient to generate vicious circularity in any possible definition of an experiential state per the constraints of this fourth picture of phenomenal consciousness.

6.2. Vicious Circularity in the Definition of Experiential States

I shall argue that, on this fourth picture, it would be *logically impossible* to define an experiential state in terms of its essential properties, K and R (defined in §6.1 *supra*), *except in a viciously circular manner*. Consider (α), a definition of an experiential state that incorporates both K and R:

(α) x is an experiential state =_{df} x bears R and the subject knows x is an experiential state

I have incorporated essential property K into (α) via the bit in the definiens that states that “the subject knows that x is an experiential state”. The problem with (α) is apparent: it is analytically circular. According to Humberstone (1997: 250), the analysis of a concept is analytically circular iff “that concept is (overtly or covertly) used in specifying the [analyzing] conditions c_1, \dots, c_n ”. Here, the definiendum, ‘experiential state’, is overtly used in the analyzing conditions in the definiens. So (α) is analytically circular.

This circularity of (α) cannot be cured by removing K from the picture. To remove K from the picture would be to abandon the picture, which is expressly built on the premise that the subject has infallible knowledge about her experiential states, entailing that she knows of any experiential state she instantiates that it is one. If the instantiation of R *sans* K were sufficient for the instantiation of an experiential state, a subject might instantiate an experiential state without her knowing that she does, which would be metaphysically impossible according to this picture. Nor can we coherently suppose that the circularity would disappear as our knowledge of phenomenal consciousness progressed and we found out all the essential properties included in R. No matter what the essential properties encompassed by R might be, K’s inclusion in the definiens of (α) would on its own be sufficient for (α)’s analytical circularity. The inclusion of K guarantees that the definiens would refer to an experiential state, which is the definiendum.

Of course, since the late 1970s, some philosophers have argued that we should not automatically classify all analytically circular definitions as vicious. The thought is that some analytically circular definitions might provide some measure of philosophical illumination to an agent who already understood the definiendum (ibid.). However, there is a type of definitional circularity, which Humberstone (1997) and Burgess (2008) call ‘inferential circularity’, which they maintain is always vicious as it can offer no degree of illumination of the definiendum. I shall show that in addition to being analytically circular, (α) is inferentially circular, and therefore viciously circular.

According to Burgess, a definition is inferentially circular iff it is “fully inferentially ungrounded”. A definition is fully inferentially ungrounded iff

the instructions for ascertaining whether or not the definiens holds [either for verifying or falsifying it] contain an unavoidable recursive loop, i.e., there is a compulsory instruction that directs us back to the starting-point [i.e., back to the definiendum] (Burgess 2008: 220).

Burgess furnishes the following example of an inferentially circular definition:

6. x is a cow =_{df} Prince Charles knows that x is a cow.

Clearly (6) is analytically circular—if we do not already possess the concept of a cow and the knowledge that this is what the English word ‘cow’ stands for, we have no hope of understanding the definiens. [...] But (6) is also inferentially circular. Suppose we are wondering whether some animal—say, Wendy—is a cow [...] Even though we probably do not possess an adequate analysis of the concept of knowledge, there seems to be near universal agreement that the truth of the propositional object of knowledge is a necessary condition for a knowledge attribution to be true. In other words, to establish that Prince Charles does know that Wendy is a cow, we have to establish that she is indeed a cow, i.e., we need to engage in a procedure that establishes (or fails to establish) the applicability of the definiendum to Wendy. (6) would therefore be useless to us, for when attempting to apply it, we receive an instruction that simply directs us back to our starting point (Burgess 2008: 221-22).

As Burgess points out, the factive verb ‘know’ in the analyzing condition in (6) makes it inferentially circular. Verifying whether Prince Charles knows that Wendy is a cow entails ascertaining whether she is, in fact, a cow. To determine whether she is a cow, we must consult the definiendum to establish its applicability to her. We have in the definiens a compulsory loop back to the definiendum, to ‘ x is a cow’, which then directs us to its definiens, which again loops us back to the definiendum, *ad infinitum*. The loop is compulsory, i.e., unavoidable, as the definition in (6) provides no other way to verify that Wendy is a cow.¹⁰

¹⁰ An anonymous reviewer suggested that we should not consider definitions fitting the pattern of Burgess’ (6) [x is a cow =_{df} Prince Charles knows that x is a cow] to be viciously circular, adverting to the fact that we use definitions of intelligence such as the Turing test, which defines machine intelligence in terms of human subjects’ judgments attributing intelligence to machines in the imitation game. However, the Turing test is not meant as a *definition* of intelligence. It is merely a *sufficiency test* and does not state any *necessary* conditions for intelligence. An intelligent but non-verbal human would fail the test.

Similarly, the factive verb ‘knows’ occurring in the definiens of (α) would require us to determine whether x is an experiential state. In the definiens of (α) , we receive an instruction requiring us to verify that the subject knows that x is an experiential state, which in turn requires us to verify that x is, *in fact*, an experiential state. This would involve a compulsory loop back to our starting point, to the definiendum, which then sends us to the definiens where we again encounter the same analyzing condition—the claim that the subject knows that x is an experiential state and so *is* an experiential state, which again sends us back to the definiendum, *ad infinitum*.¹¹ Our definition of an experiential state, (α) , is both analytically and inferentially circular, just as Burgess’s (6). Therefore, it is viciously circular.

A phenomenal realist cannot remedy the situation by supposing that the subject has a mere *true belief*, but lacks *knowledge* about her experiential state (perhaps on the supposition that the subject would lack adequate justification for her true belief and so would not count as possessing knowledge), and so amend the definition in (α) to $(\alpha)^{TB}$:

$(\alpha)^{TB}$ x is an experiential state $=_{df}$ x bears R and the subject **truly believes** that x is an experiential state.

The problem is that ‘true belief’ is a factive state, just as knowledge is. True belief entails that the proposition believed is true. In the definiens of $(\alpha)^{TB}$, we receive an instruction requiring us to verify that the subject truly believes that x is an experiential state, which in turn requires us to verify that x is, *in fact*, an experiential state. Otherwise, the subject’s belief would be false. Verifying that x is an experiential state would involve a compulsory loop back to our starting point, to the definiendum, which then sends us to the definiens where we again encounter the same analyzing condition—the claim that the subject truly believes that x is an experiential state and so *is* an experiential state, which again sends us back to the definiendum, *ad infinitum*.

And we cannot remedy the problem by amending the definition of $(\alpha)^{TB}$ by changing *true belief* to *mere belief* and amending our definition of experiential states to $(\alpha)^B$, as follows:

$(\alpha)^B$ x is an experiential state $=_{df}$ x bears R and the subject **believes** that x is an experiential state

This amendment would fail because it is too weak to capture what this fourth picture posits—that the subject’s belief is necessarily true. It leaves open the possibility that the subject’s belief might be false. By *reductio*, under such a definition of an experiential state, if a p-zombie believed (erroneously) that x was an experiential state, x would be an experiential state and the p-zombie would be phenomenally conscious.

Because any definition of an experiential state would necessarily be viciously circular—given that according to this picture property K would need to be included in any definition of experiential states, and that would be sufficient to

¹¹ The applicability of the definiens of (α) to x would be *falsifiable* by establishing that x fails to bear R. However, (α) would nevertheless be inferentially circular. According to Burgess, it is sufficient for inferential circularity that the instructions for ascertaining whether or not the definiens holds—*either* for verifying *or* falsifying it—contains an unavoidable recursive loop. (α) ’s compulsory recursive loop in verifying whether the definiens holds would therefore be sufficient for (α) ’s inferential circularity.

generate the vicious circularity—I conclude that this fourth picture of phenomenal consciousness is incoherent. To be clear, my claim is not merely that a *particular proposal* for a definition of experiential states is viciously circular. I’m claiming that it would be *logically impossible* to define experiential states but in a viciously circular manner under the constraints of this fourth picture of phenomenal consciousness. This entails that the conditions for the instantiation of an experiential state would have to be circular conditions. Unless it’s logically possible for a set of circular conditions to obtain (which it’s not), it’s logically impossible for experiential states to be instantiated according to the constraints of this fourth picture of phenomenal consciousness.

7. Conclusion

I have argued that it would be impossible for phenomenal consciousness to exist without both the experiential and the cognitive state being included in the picture, for we cannot make sense of either of the one-component pictures I discussed in §4 and §5. At the same time, we cannot make sense of any picture that includes both experiential and cognitive states, for both two-component pictures fail, as I argued in §3 and §6. None of the four pictures is coherent.

Illusionism offers a two-component picture but without incoherence. The illusionist would substitute *sensory states* for experiential states, and our cognitive states would be about them instead. Unlike experiential states, sensory states would *be functional, neural, or other sorts of physically grounded states* that could in principle be fully third-personally scientifically characterized and defined in functional/structural/neural terms without regard to whether any subject was aware of them or knew what they were like. Unlike experiential states, which would be necessarily conscious (as I argued in §5), sensory states could be either conscious or unconscious. Our awareness and knowledge of our sensory states would *not* be a necessary condition for their instantiation, as was the case for experiential states. Illusionism thus avoids the problematic intertwining of the cognitive and the experiential states we encountered in both two-component pictures of phenomenal consciousness. Furthermore, illusionism is an attractive option because it would preserve the physicalist picture, which is the empirically best-supported theory of the underlying metaphysical substrate of the mind.

An anonymous reviewer objected that I have not ruled out the mysterian position, which posits that we humans are cognitively too limited to understand phenomenal consciousness (McGinn 1989; Nagel 1989). Any solution to the hard problem of consciousness might lie outside the bounds of our cognitive capacity to discover or understand it. Understanding phenomenal consciousness would be “cognitively closed” to us (McGinn 1989: 350). I shall agree tentatively with this objection. That’s why I stated in the introduction that I believe illusionism is only *probably* true. But I shall argue that while mysterianism cannot be ruled out, it is far less plausible than illusionism.

I take it that I have at least shown that *our* concept of phenomenal consciousness—which is constituted by (or at least heavily informed by) our intuitions about it, is incoherent. So we ought to reject *that* conception of phenomenal consciousness. Would that be a good reason to reject the concept of phenomenal consciousness *altogether*, as the illusionist would propose? Let’s suppose there *was* some coherent explanation of phenomenal consciousness somewhere out there, maybe in the minds of cognitively superior aliens from another world. Assume

that these aliens are phenomenally conscious in the same way we are; they are just much more intelligent than we are. These aliens' conception of phenomenal consciousness would surely be a *substantially revised* conception, given the incoherence of our conception. Now, what criteria could we use to tell that the aliens' revised conception was about the *same explanandum*, the thing we call 'phenomenal consciousness'? What would make the aliens' explanation count as an explanation of phenomenal consciousness rather than some other phenomenon? If we could somehow manage to understand the aliens' explanation, I doubt we'd consider it a good explanation, for it would gainsay our core intuitions about phenomenal consciousness. We might characterize their explanation as *explaining away* phenomenal consciousness rather than explaining it, as certain philosophers did with Daniel Dennett's brilliant 1991 *Consciousness Explained*. The aliens' explanation of phenomenal consciousness might even turn out to *be* illusionism. How could any 'phenomenal residue' (Frankish 2012: 669) remain in these aliens' explanation when it has stripped our core intuitions away? It's not as if we can specify the explanandum in some way other than by appealing to our raw intuitions supported by the usual thought experiments that pump them and make them vivid (p-zombies, Mary the color scientist, the inverted spectrum, et al). We cannot point at phenomenal consciousness, and we cannot locate it in space or in the brain. Even words fail us when we try to explain what it is, and we always end up resorting to thought experiments. Our first-person intuitions are all we have, and they make no sense. I believe any conceptual revision would have to be so radical that it would constitute changing the subject, such that the revised conception would be unrecognizable to us as being about the same explanandum. That is why I find mysterianism to be implausible, although it cannot be ruled out altogether.

References

- Burgess, J.A., 2008. When is circularity in definitions benign?. *The philosophical quarterly*, 58 (231), 214–33.
- Chalmers, D.J., 1996. *The conscious mind: in search of a fundamental theory*. New York: Oxford University Press.
- Chalmers, D.J., 2003. The content and epistemology of phenomenal belief. In: A. Jokic and Q. Smith, eds. *Consciousness: New philosophical perspectives*. New York: Oxford University Press, 220–271.
- Chalmers, D.J., 2006. Phenomenal concepts and the explanatory gap. In: T. Alter and S. Walter, eds. *Phenomenal concepts and phenomenal knowledge: New essays on consciousness and physicalism*. New York: Oxford University Press, 167–94.
- Chalmers, D.J., 2018. The meta-problem of consciousness. *Journal of consciousness studies*, 25 (9-10), 6–61.
- Dennett, D.C., 1991. *Consciousness explained*. (P. Weiner, Illustrator). Little, Brown and Co.
- Dennett, D.C., 1993. *Consciousness explained*. London: Penguin.
- Dennett, D.C., 2006. *Sweet dreams: Philosophical obstacles to a science of consciousness*. Cambridge, MA: MIT press.

- Duncan, M., 2023. How you know you're conscious: Illusionism and knowledge of things. *Review of philosophy and psychology*, 14 (1), 185–205.
- Frankish, K., 2012. Quining diet qualia. *Consciousness and cognition*, 21 (2), 667–76.
- Frankish, K., 2016. Illusionism as a theory of consciousness. *Journal of consciousness studies*, 23 (11-12), 11–39.
- Frankish, K., 2021a, on *Hear this idea* (podcast), hosts: Fin Moorhouse and Luca Righetti, episode thirty-nine, “Keith Frankish on illusionism about consciousness” 22 Nov. 2021, <https://hearthisidea.com/episodes/keith> [Accessed 17/05/2024].
- Frankish, K., 2021b. Panpsychism and the depsychologization of consciousness. *Aristotelian Society Supplementary Volume*, 95 (1), 51–70.
- Giustina, A., 2021. Introspection without judgment. *Erkenntnis*, 86 (2), 407–27.
- Goff, P., 2020. on *Consciousness live!* (podcast), host: Richard Brown, S2 E7, “Discussion with Keith Frankish and Philip Goff”. https://www.youtube.com/watch?v=3zqhOETKJ6U&ab_channel=RichardBrown [Accessed 17/05/2024].
- Horgan, T., 2012. Introspection and phenomenal consciousness: Running the gamut from infallibility to impotence. In: D. Smithies and D. Stoljar, eds. *Introspection and consciousness*. Oxford: Oxford University Press, 405–22.
- Humberstone, I.L., 1997. Two types of circularity. *Philosophy and phenomenological research*, 57 (2), 249–80.
- Jackson, F., 1982. Epiphenomenal qualia. *Philosophical quarterly*, 32, 127–36.
- Kripke, S., 1980. *Naming and necessity*, Cambridge, MA: Harvard University Press.
- Levine, J., 2001. *Purple haze: The puzzle of consciousness*. New York: Oxford University Press.
- McGinn, C., 1989. Can we solve the mind-body problem?. *Mind*, 98 (391), 349–66.
- Moran, R., 2001. *Authority and estrangement: An essay on self-knowledge*. New Jersey: Princeton University Press.
- Nagel, T., 1974. What is it like to be a bat?. *Philosophical review*, 83, 435–50.
- Nagel, T., 1989. *The view from nowhere*. New York: Oxford University Press.
- Russell, B., 1910-11. Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian society*, 11, 108–28.
- Schneider, S., 2019. *Artificial you*. Princeton: Princeton University Press.
- Searle, J., 1998. How to study consciousness scientifically. *Brain research reviews*, 26 (2/3), 379–87.
- Strawson, G., 2018. *The consciousness deniers*. The New York review of books, <https://www.nybooks.com/online/2018/03/13/the-consciousness-deniers/> [Accessed 17/05/2024].