



Dualism all the way down: why there is no paradox of phenomenal judgment

Helen Yetter-Chappell¹

Received: 6 September 2021 / Accepted: 9 March 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Epiphenomenalist dualists hold that certain physical states give rise to non-physical conscious experiences, but that these non-physical experiences are themselves causally inefficacious. Among the most pressing challenges facing epiphenomenalists is the so-called “paradox of phenomenal judgment”, which challenges epiphenomenalism’s ability to account for our knowledge of our own conscious experiences. According to this objection, we lack knowledge of the very thing that epiphenomenalists take physicalists to be unable to explain. By developing an epiphenomenalist theory of subjects and mental states, this paper argues that there is nothing paradoxical or problematic about the epiphenomenalist’s understanding of phenomenal judgments or phenomenal self-knowledge. The appearance of paradox emerges from inconsistently combining (epiphenomenalist) dualism about qualia with a physicalistic conception of subjects of experience. The lesson we should take from this is not that there is anything wrong with epiphenomenalism, but that epiphenomenalist dualists should be “dualists all the way down”—embracing a picture of mind that gives phenomenology a central place, in its understanding of both subjects and their knowledge of their own minds. Epiphenomenalist-friendly accounts of reference and memory are also developed, showing that neither of these issues creates a paradox for the epiphenomenalist.

Keywords Epiphenomenalism · Paradox of phenomenal judgments · Dualism · Phenomenal judgment

Epiphenomenalist dualists hold that certain physical states give rise to non-physical conscious experiences, but that these non-physical experiences are themselves causally inefficacious. Your brain state, after a mosquito bites you, causes you to have an experience of itchiness. But this itchy experience is not what’s responsible for your subsequent scratching behavior or your vocalizations cursing mosquitoes. Rather, it

✉ Helen Yetter-Chappell
hoyetter@gmail.com

¹ Department of Philosophy, University of Miami, Coral Gables, FL, USA

is the brain state sending signals to your arms and vocal cords that is responsible for this behavior. Among the most pressing challenges facing epiphenomenalists is an argument purporting to show that epiphenomenalism is incompatible with our having knowledge of our own conscious experiences. If this argument is right, we lack knowledge of the very thing that epiphenomenalists take physicalists to be unable to explain. This “paradox of phenomenal judgment” has been discussed at length by philosophers including Shoemaker (1975), Chalmers (1996, 2003), Kirk (2005, 2008), Pauen (2006), and Robinson (2006, 2012).

This paper argues that there is nothing paradoxical or problematic about the epiphenomenalist’s understanding of phenomenal judgments or phenomenal self-knowledge. The appearance of paradox emerges from inconsistently combining (epiphenomenalist) dualism about qualia with a physicalistic conception of subjects of experience. The lesson we should take from this is not that there is anything wrong with epiphenomenalism, but that epiphenomenalist dualists should be “dualists all the way down”—embracing a picture of mind that gives phenomenology a central place, in its understanding of subjects, their mental states, and their knowledge of their own minds.

The structure of the paper is as follows. In Sects. 1–6, I respond to the charge that epiphenomenalism is incompatible with knowledge of qualia. I develop an epiphenomenalist theory of mind, considering what epiphenomenalists should believe about the nature of subjects (Sect. 2), beliefs (Sect. 5), and knowledge (Sect. 6). I put this to work illustrating the overall coherence of the epiphenomenalist worldview. Sect. 7 considers two related challenges to epiphenomenalism, which concern the epiphenomenalist’s ability to account for phenomenal memory and reference to qualia. I show how the epiphenomenalist can meet these challenges, and why none prove paradoxical.

1 The alleged paradox

In its most basic form, the paradox takes the form of a *reductio* on epiphenomenalism: If epiphenomenalism is true, we can’t have knowledge of our own conscious experiences. But we can have such knowledge. So epiphenomenalism is false. Fleshed out¹:

1. If epiphenomenalism is true, I have a zombie twin (z-twin) in another possible world.
2. As she scratches at her leg, the very same things go on inside of her brain as inside of my brain. These brain processes cause the very same sounds “Itchiness feels like *this*” to come out of her mouth as come out of my mouth.
3. So my z-twin has the very same phenomenal belief as me—“Itchiness feels like *this*”—formed by the very same mechanism.
4. But my z-twin’s belief is not only false, it’s not justified.
5. If her belief is not justified, and my belief was formed by the same mechanism, then my beliefs can’t be justified either.
6. So—if epiphenomenalism is true—my phenomenal belief isn’t justified.
7. But it clearly is justified.

¹ This follows Chalmers’s (2003) presentation of the paradox.

8. So epiphenomenalism is false.

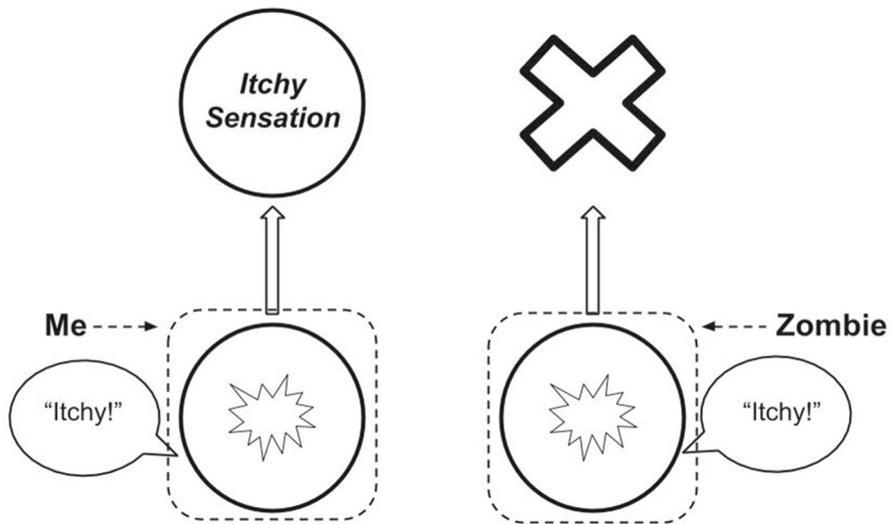
Chalmers (1996, 2003) has offered two attractive responses to this paradox: First, one might dispute the notion of justification presupposed by premise (5) (Chalmers, 1996). Second, one might reject premise (3) on grounds that zombies either (a) lack beliefs, or (b) lack *phenomenal* beliefs. As Chalmers (2003) puts it,

The basis of intentionality is poorly understood, and one might plausibly hold that a capacity for consciousness is required for intentional states. But even if we allow that zombies have beliefs, it is clear that a zombie cannot share a conscious being's phenomenal beliefs. The content of a conscious being's direct phenomenal beliefs is partly constituted by underlying phenomenal qualities. A zombie lacks those qualities, so it cannot have a phenomenal belief with the same content.

While I agree with Chalmers that rejecting premise (3) is an attractive response and the right way to go, it is not (by itself) sufficient to dissolve the challenge the epiphenomenalist faces. As Kirk (2008) notes, this response presupposes that we stand in a certain epistemic relationship to our qualia such that they can be constituents of our thoughts. (Not just any quality and not just any qualia can be a constituent of my thoughts—I must stand in the right relation to the qualia if this is to be possible.) And, as Kirk argues, there's reason to think that the epiphenomenalist cannot accommodate such a relationship.

This paper will diagnose the original paradox as stemming from inconsistently trying to combine epiphenomenalist *dualism* about qualia with a *physicalistic* conception of subjects of experience. Once we appreciate the dualist's view of subjects as inherently phenomenal, a rejection of premise (3) naturally falls out. But furthermore, a solution to Kirk's development of the paradox naturally presents itself—for Kirk's argument stems from inconsistently trying to combine epiphenomenalist *dualism* with a *physicalistic* conception of mental states. As before, an appreciation of the dualist's view of mental states as inherently phenomenal undermines Kirk's objection. In both instances, the paradox involves taking a piece from the dualist's worldview and a piece from the physicalist's, and alleging a problem when they don't fit together. But we never should have expected those two views to cohere. There is only a problem for epiphenomenalism if the *epiphenomenalist dualist pieces* fail to form an adequate picture of our minds and our judgments.

Let's begin with the initially stated challenge to beliefs about qualia. In essence, the challenge raised by the "paradoxes" is something like this: I am a physical creature; my beliefs are physical states. The epiphenomenalist tells us that qualia are non-physical, and that they don't have any causal impact on the physical world. It follows that they don't have any impact on me or on my beliefs. So they are irrelevant to me and my beliefs. So I could not possibly latch onto them to think about them or have any knowledge of them. But this is clearly mistaken; so epiphenomenalism is false. (Below is an illustration of the described disconnect between me/my beliefs and my experiences, in relation to my z-twin.)

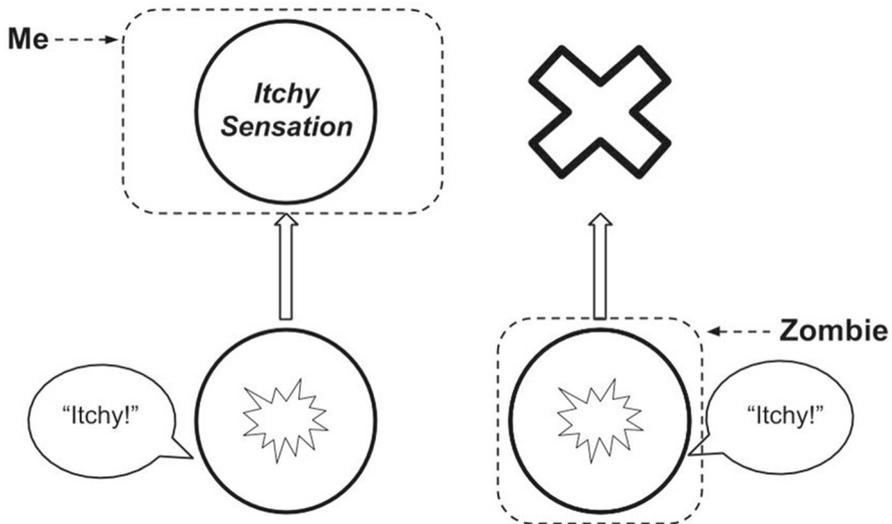


2 The basic diagnosis

Given the framing above, it's easy to see that the paradox implicitly assumes a physicalistic conception of belief and of subjects—a conception that should be anathema to dualists. The dualist should insist that my conscious experiences are essential to *me* and to my *beliefs*: comprising (or partially comprising) both me and my phenomenal beliefs. Chalmers (1996) makes a similar point in passing.² But the nature of the subject and the relation between subject and qualia don't play a central role in Chalmers's treatment of the paradox. I'll argue, however, that the nature of subjects and their mental states is crucial to understanding this cluster of challenges to epiphenomenalism. It is only when we fail to acknowledge the essential connection between subjects and their qualia—when we mismatch a physicalistic conception of subjects and their mental processes with a dualistic conception of consciousness—that consciousness becomes irrelevant. The problem here is not the picture of consciousness; rather it's the picture of the rest of the mind that renders consciousness disconnected from it. The moral that we should take away from this is that dualists should be dualists all the way down: If qualia are non-physical, subjects and their mental states are (at least in part) non-physical too. Below is an illustration of how the thoroughgoing dualist understands the relationship between subjects and their conscious experiences, in contrast to their zombie counterparts.³

² “[T]he question is whether *I* am justified in the belief, not whether my *brain* is justified in the belief, and if property dualism is correct then there is more to me than my brain. I am constituted by both physical and non-physical properties, and the full story about me cannot be told by focusing only on one half.” (Chalmers 1996, p. 198).

³ The illustration treats subjects as solely constituted by non-physical qualia. The brain is responsible for generating the subject and their dispositions, but is not itself a part of the subject. Alternatively, the



This naturally entails a rejection of premise (3) of the initial paradox. This premise made two claims: (i) my z-twin has the same phenomenal beliefs as me—after all, the same things happen in her brain as mine, causing the same sounds to leave her lips. And (ii) her belief is formed by the same mechanism as mine—after all, our brains are doing precisely the same things. Arguably, given the correct (dualist) understanding of subjects and their beliefs, neither of these is true.

(i): If consciousness is essential to the phenomenal beliefs I have, then my z-twin does not have the same phenomenal beliefs as me, since she doesn't have any such experiences. My z-twin's brain and vocal cords may do the same things as mine, but unless we embrace a physicalistic conception of phenomenal beliefs that renders them incapable of coming apart from these physical processes, there is no reason to think that my z-twin's phenomenal beliefs must also be the same as mine. And the dualist has no reason to embrace such a physicalistic conception of phenomenal beliefs. (Chalmers, 2003) (More on this to follow in Sect. 3.)

(ii): There is likewise no reason for dualists to accept the claim that our beliefs were formed by the same *mechanisms*. If the sole mechanism at work in forming my beliefs was the physical workings of my brain, it would follow that my z-twin's beliefs are formed by the very same mechanism. But—insofar as the dualist holds that (non-physical) phenomenal experiences are essential constituents of phenomenal beliefs—my phenomenal beliefs are formed not only by the physical workings

Footnote 3 continued

dualist might hold subjects to be *partially* constituted by non-physical qualia and partially constituted by the physical brains that generate these experiences. In the interest of simplicity, I will write in terms of the former view, with differences noted in footnotes. Epistemic contact is a mental phenomenon. The mental part of me—be that me in my entirety or merely a proper part of me—is what's essential to my epistemic states. Thus, the same response to the paradox will be available regardless of which view of subjects the dualist embraces. Call the mental part of the picture above *m*, and the physical part *p*. We might identify the subject with *m*, and *p* as the mere physical base of the subject—in which case, subjects (simpliciter) have epistemic contact. Or we might identify the subject with *m + p*—in which case, subjects have epistemic contact thanks to their *m*-parts.

of my brain, but by the bridging laws responsible for generating these phenomenal experiences. No such mechanism is at work for my z-twin.

3 Phenomenal judgments

Not only are dualists not required to embrace (i) and (ii), they have well-motivated reasons to reject each, insofar as they take consciousness to have a central place in our minds, as something integral to us, our thoughts, and our epistemic positions.

The popular constitutional theory of phenomenal concepts (Balog, 2012; Chalmers, 2003; Papineau, 2002) offers dualists a compelling way of fleshing out this central role for consciousness: rendering phenomenology a constituent of our phenomenal concepts, and (hence) our phenomenal judgments. If phenomenal experiences are fundamentally non-physical, the constitutional theory entails that phenomenal concepts (and the judgments that employ these concepts) are also, at least in part, non-physical. Chalmers (2003) argues that a constitutional theory of phenomenal concepts can undermine the alleged paradox, by developing a systematic basis for taking your zombie twin's judgments to differ from your own.

The rough idea behind the constitutional theory is that, much as I might *use* a song to think about the song itself (“I wish I remembered the name of that song that went *bum bum bum bum bu-dum*”), so I can think about my phenomenal experiences by using instances of the relevant type (“I hate *itchy sensation*”). When I think this, I am directly acquainted with the phenomenal experience of itchiness, for the experience is—quite literally—a part of my thoughts. By contrast, my z-twin does not have any phenomenal experiences, and hence does not share either my phenomenal concepts or my phenomenal beliefs.

The constitutional theory of phenomenal concepts also gives us a way of understanding what it takes for my phenomenal beliefs to be justified that doesn't rely on a causal connection between my beliefs and their objects. When I believe, e.g., “Itchiness feels like this”, my thought is itself composed from the objects of my belief. I do not stand “at a distance” from its truth-makers, as with a causal relationship. Rather, the nature of the object of my belief—the essence of itchiness—is directly presented to me (it is a constituent of me) as the means by which I think about itchiness. Because the very phenomenal experience that is the object of my thought is being used to do the thinking, the truth of my judgment is even more secure than in the ordinary case.⁴

While the constitutional theory of phenomenal concepts has many theory-neutral virtues (Balog, 2012), it is particularly attractive for dualists as it gives consciousness a central place in our minds: not just as a tacked-on extra, but as something that is integral to us, our thoughts, and our epistemic position. (But note that what's crucial to the rejection of (i) is simply that consciousness be essential to our phenomenal judgments. The constitutional theory gives a compelling basis for holding this—one that I'm partial to—but there may be other ways of making sense of this central place for phenomenology (e.g. Pitt, 2011).)

⁴ For developed accounts of relationship between the constitutional theory of phenomenal concepts and the direct acquaintance see: Chalmers (2003) and Gertler (2011, 2012).

Thus far, the response I've sketched to the alleged paradox is fairly familiar—mirroring treatment by Chalmers (2003). But while this provides us with a promising start for addressing the (putative) paradox, it is not sufficient to dispel the mystery. We're not epistemically acquainted with just any old qualia. Consider: My experience of pain or itchiness might be able to serve as a constituent of my thoughts, but *your* experience of pain cannot be a constituent of my thoughts, any more than a teacup or a basketball can. You simply cannot stand in the relevant epistemic relationship of direct acquaintance with my thoughts or with a basketball. Similarly, spontaneously introducing a red quale into a zombie world wouldn't necessarily give any of the zombie inhabitants epistemic access to this red quale (cf. Kirk, 2008, p. 77). So for this defense to have teeth, we need an explanation as to *how it is that you can stand in the relevant epistemic relationship to your experiences, such that they can be constituents of your thoughts*.

Kirk (2005, 2008) develops this point with an ingenious example, designed to show that no such explanation can be given. Without an answer to Kirk's challenge, while we may have identified a flaw with the original paradox, we have not shown that epiphenomenalists can account for the epistemic contact that we clearly have with our own qualia. I'll argue that, as with the initial paradox, Kirk's argument relies on a physicalistic conception of subjects and their mental states, which the dualist has no reason to accept. The explanation of how it is that we stand in the relevant epistemic relationship to our experiences emerges from the correct account of the relationship between subjects, their mental states, and their qualia.

4 Kirk's Currents

Kirk (2008, p. 78) writes:

(Q) How, in an E-world,⁵ could physical processes in an individual body contribute to anyone's being in the relevant sort of epistemic contact with e-qualia: what could hook them up, epistemically?

I will argue that e-qualists cannot give a satisfactory answer to this question. The structure of an E-world would prevent its inhabitants from being in epistemic contact with 'their' e-qualia because the cognitive processing essential for the relevant sort of epistemic contact would have to be performed by physical processes epistemically insulated from all e-qualia.

Epiphenomenalists can't hold that we're epistemically acquainted with just any old qualia: I'm not epistemically acquainted with your qualia; if there were a zombie world where a red quale was suddenly introduced to the world, none of the zombies would magically gain epistemic contact with it. To avoid these problems, Kirk suggests that to have epistemic contact with epiphenomenal qualia, the qualia must be "caused by and isomorphic to the relevant physical processes" (2008, p. 78).

But now the epiphenomenalist faces a serious difficulty: Suppose that epiphenomenalism is true. Now imagine that in addition to producing qualia, our brain processes

⁵ Let an 'e-world' be a world in which epiphenomenalism is true, and 'e-qualia' be epiphenomenal qualia.

(the very same brain processes) also produce “minute patterns of electrical activity which are in relevant respects isomorphic to them—but have no effects on them” (2008, p. 79). We never observe these currents nor do we ever come to know anything about them. Call these “Kirk’s Currents” (KC). It certainly seems possible that there be such currents. But we now have an example of something “caused by and isomorphic to the relevant physical processes”, but where we are *not thereby in epistemic contact with it*. So the same criterion for epistemic contact cannot be sufficient for epistemic contact with epiphenomenal qualia.

It follows that something more than causation plus isomorphism must be necessary for epistemic contact. Kirk sees three options for explaining the distinctive contact we have with our qualia.

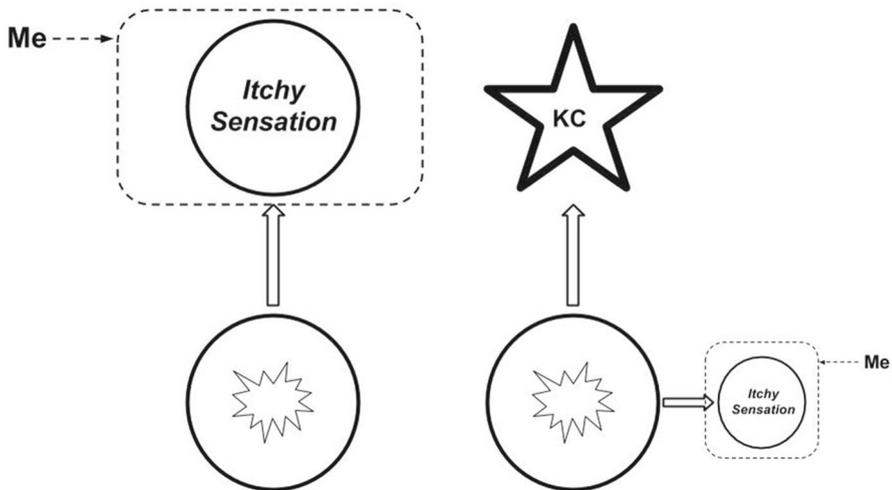
1. Intrinsic properties of the physical relata (our brains)
2. Intrinsic properties of qualia
3. The distinctive relation between our brains and qualia

But, Kirk argues, it’s hard to see how any of these could adequately distinguish e-qualia from Kirk’s Currents: (1) The very same physical components are involved in generating both e-qualia and KC, so this cannot explain how I could have epistemic contact with one and not the other. (2) KC’s “intrinsic properties could be whatever you please (provided that they remain inert); I should still not be able to notice, think about, attend to, remember, or compare them” (2008, p. 79). So it’s hard to see how differences between these inert intrinsic properties could facilitate epistemic contact in the case of qualia. (3) And it’s mysterious what distinctive relation there could be between the physical properties and qualia that couldn’t be replicated for KC. It is not enough to simply say that we’re acquainted with our qualia or that they (partially) constitute our phenomenal judgments. A complete answer to the paradox must explain *how* it is that we can have this kind of epistemic contact with epiphenomenal qualia, given that we cannot have such contact with other epiphenomena.

5 Resolving the paradox

This analysis of the relationship between e-qualia and Kirk’s Currents follows naturally from a physicalistic conception of subjects and their beliefs. But—as in the reply to the initial paradox—there’s no reason for a dualist to accept such conceptions. It’s only when presupposing that qualia are something separate from me (like KC) that the two cases are parallel. If e-qualia are a part of me, I stand in a different relation to my (epiphenomenal) qualia than to Kirk’s Currents. The distinctive contact I have with my qualia is thus explained by a fourth option that Kirk overlooks: the distinctive relation between *me*, qua subject of experience, and my qualia.

The dualist will model the difference between qualia and Kirk’s Currents as illustrated below.



Kirk argues that even so, if qualia are epiphenomenal, they are “epistemically insulated from all cognitive processes. No subject could think about, notice, attend to, or remember items in that stream of consciousness: no one could be in epistemic contact with them in the relevant sense” (2008, p. 81). Regardless of whether we individuate subjects so that qualia are part of them or not, if these qualia are causally inert, they will not affect our beliefs; they will play no role in our noticing our itches or attending to our pains. In effect, the argument seems to be: E-qualia may (partially) constitute the subject, but they are cut off from the cognitive processes that are essential for epistemic contact.

This conception of epistemic contact relies on a presumption that causal interaction is the *only* way we can have epistemic contact with (e-)qualia. This is a natural assumption for a physicalist to make, reasoning as follows:

- i. To have epistemic contact with a non-abstract entity, our mental states must either causally interact with it or contain it.
- ii. (E-)qualia are separate from us, and hence are not literally contained within our mental states.
- iii. So to have epistemic contact with (e-)qualia, they must causally interact with our mental states—i.e. there must be cognitive processes that causally link our mental states with the relevant qualia.

As we’ve seen, dualists take qualia to be essential to subjects. Further, I’ll argue that dualists can and should take qualia to be essential to our mental states. There is no motivation for a dualist to accept (ii).

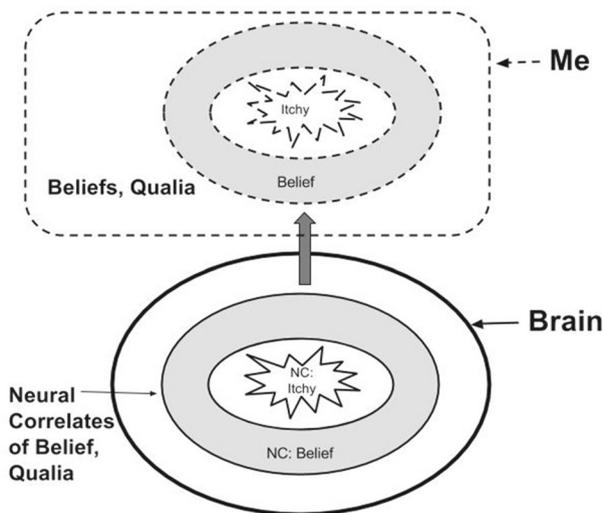
Much as the initial puzzle generated tension by combining a physicalist conception of subjects with a dualist conception of consciousness, this seems to require a physicalistic (or, at very least, an anti-epiphenomenalist⁶) conception of epistemic

⁶ It would be open to an interactionist to hold that qualia are integral to us and our mental states *only if* they causally interact with brains. This would render e-qualia separate from us. But while this is logically

contact—and hence not one that can be used as a starting point in arguing that epiphenomenalism is incoherent. As before, the dualist need not (and should not) accept this.

It was supposed that qualia are isomorphic to brain processes. A natural extension of this would hold that all of the occurrent thoughts, desires, noticings, etc. that my brain generates have isomorphic phenomenal parallels. (The dualist should expect anything we're conscious of—including the fact that I'm currently *desiring* a cup of tea, rather than *remembering* one—to be manifest in the “phenomenal realm”).

On this view, the mental states—beliefs, thinkings, attendings, and so on—that Kirk takes to afford us with epistemic contact are not (purely) physical states, but are (at least partially)⁷ phenomenal ones:



Thus my z-twin's lack of phenomenology does not merely result in her having different *phenomenal* beliefs from me (virtue of her failing to have phenomenology to partially constitute her phenomenal beliefs). My z-twin lacks any beliefs whatsoever (although she has the neural underpinnings of beliefs). Because of this we cannot (as Chalmers, 1996 proposes) frame the paradox in purely psychological/functional

Footnote 6 continued

consistent with interactionism, it's unmotivated for a dualist—for unless you implicitly identify subjects with their brains, what motivation could there be for thinking that qualia can be part of *me* only if they interact with my *brain*?

⁷ As for subjects (ftn. 3), there might be broader and narrower versions of mental states: the narrower one holding that mental states are purely phenomenal processes that our ghost-twins could share; the broader ones holding that mental states have both physical and phenomenal aspects. I think it's clear that my ghost twin and I could share the same beliefs, noticings, etc., and so favor the narrower version, though the defense of epiphenomenalism does not rely on this. As before, we might call the conscious part of the above picture *m* and the physical part *p*. Whether we identify mental states with *m* alone, or with *m + p*, the suggestion is that it is the *m* part which is relevant to epistemic contact. While on the narrow view, epistemic contact is the result of a distinctive relation between *me/my* mental states *in their entirety* and my qualia, on the broader view, it will be a result of a distinctive relation between *the conscious part of me/my* mental states and my qualia.

terms. If we restrict ourselves to considering “what I and my zombie twin have in common” (Chalmers, 1996, p. 174), we will omit epistemic contact with qualia from the picture. This is not because epiphenomenalists are unable to account for such contact, but because we will have eliminated from consideration the very thing that dualists should think makes such contact possible.

So this is the picture of the mind we have arrived at: My mental states and my sensations are all parts of *me*, present within my mind. When I notice that my leg feels itchy, the itchiness that I feel partially constitutes my noticing. And this noticing—along with my belief that a mosquito caused the sensation, my desire that the sensation cease—are all constituents of *me*. On this nested phenomenal picture, there is no distance between experiences, the mental states that involve them, and the subjects who have them. The relationship between these things is not causal; it is constitutive.

I (as a conscious, thinking, subject of experience) may not have epistemic access to the neural processes underwriting my mental states. But I do have access to the mental states themselves, provided these are understood in accordance with dualism: as phenomenal. Thus my epistemic contact with e-qualia is not understood as the result of a “distinctive relation between physical properties and qualia”. Rather, this epistemic contact is the result of a distinctive relation between *me/my mental states* and my qualia. As in the response to the initial paradox, the crucial move concerns the epiphenomenalist’s understanding of subjects and their mental states. The epiphenomenalist should embrace an essential connection between subjects and experience.

This gives a positive epiphenomenalist-friendly explanation of our epistemic contact with e-qualia. But to fully address Kirk’s challenge, we need to understand why the situation is not analogous for contact with Kirk’s Currents. Why can’t subjects and their mental states be partially constituted by KC in a parallel way? The answer lies in the intrinsic nature of subjects (and their mental states) and their qualia. Qualia aren’t just any old non-physical property, nor are they just any old epiphenomenal property (cf. Chappell, 2019). Qualia (experiences) are essentially *experienced*. And for something to be experienced there must be an *experiencer*.⁸ By contrast, there is no basis for positing an essential connection between KC and subjects of experience.⁹ In short, qualia, but not Kirk’s Currents are essentially parts of subjects of experience. And it is this constitutive connection to the subject of experience that affords us epistemic access to qualia.

5.1 Subjects and qualia

Embracing an essential connection between experiences and subjects doesn’t in itself tell us what subjects *are*, even given dualism as a background assumption. One natural picture would have it that subjects are something over and above—ontologically distinct from—the experiences that they experience. By contrast, Strawson (2003)

⁸ An essential connection between qualia and subjects has been endorsed by many philosophers, including Descartes (1641), Frege (1918), Strawson (2003, 2017), and Zimmerman (2011), and is a natural position for a dualist to embrace.

⁹ Robinson (2016, p. 55) offers a closely related objection to Kirk, noting that “[t]here is nothing ... to make the cranial currents the *intentional object* of the epistemic states, whereas there will be reasons to take the phenomenal content as the intentional object”.

argues subjects might be nothing over and above the experiences themselves—provided that we properly understand experiences as things that are *by their very natures experienced*. I take it either option is available to the epiphenomenalist.

It might be objected that this leaves unaddressed the question of *why* an experience (e.g. a red quale) belongs to a given subject—and hence fails to fully address the question of how I can have epistemic contact with my red quale. Mustn't there be some additional fact that makes the quale belong to me (rather than to you or no one)?

Consider an analogy. Suppose we are asked what makes it the case that a particular tree is in location L (on Loren's farm), rather than in location M (on Marian's). The answer will vary depending on your view of the nature of space. But whether you're a substantivalist or a relationalist, the answers on offer are apt to seem trivial. A substantivalist might hold that the farm occupies a particular region of absolute space. The tree is on the farm because it has the property of being located at the absolute spatial coordinates L, within this region. For the relationalist, the tree's location on the farm is instead constitutively explained by the tree being co-located with certain other objects (namely, those of Loren's farm). No reference to absolute space is required.

What makes it the case that my red quale is 'located' in my mind, i.e. that it's *my* quale? The candidate answers will vary depending on whether a 'thick' or 'thin' view of subjects is embraced—but, again, the answers are similarly trivial. If the epiphenomenalist embraces a 'thick' view of subjects as distinct from the experiences they have, their answer will be analogous to the substantivalist's view of location above: the quale belongs to a particular subject because it is contained within that subject's substantial mind or Cartesian soul. If the epiphenomenalist embraces a thin Strawsonian conception of subjects, their answer will be analogous to the relationalist conception above: the quale's 'location' will be explained by its being co-conscious with certain other qualia (namely, those comprising my mind). On this thin view, experience is essentially experience-for. But *who* it is essentially experience-for just is the phenomenal unity of which it is a part.¹⁰

Further challenges remain. Rudd (2000, p. 59) objects that "it is hard to see how one could construct ... a [phenomenal] unity out of discrete conscious states, caused by distinct brain processes and with no direct causal connections existing between them."¹¹ The worry is that without (i) a single locus of consciousness in the brain, or (ii) causal relations between my red quale and the other qualia comprising my mind, there's no way for them to *become* phenomenally unified.

Rather than appealing to either (i) or (ii), the epiphenomenalist might propose that the phenomenological "atoms" collectively ground their own unity. On such a view, phenomenology is generated by the bridging laws in a piecemeal fashion (e.g. yielding disunified states of phenomenal redness, bird-shapedness, and tweeting), and these discrete phenomenal components ground our unified phenomenology. Whether

¹⁰ A thick view of subjects would naturally go hand in hand with substance dualism. On a thick view of subjects, the subject can persist independently of its phenomenal properties, and hence must be something distinct from them. This combined with the *subject* dualism that I've argued epiphenomenalists must embrace would arguably commit us to mental substances. By contrast, a thin view of subjects does not (in itself) entail any commitment to substances at all.

¹¹ Thanks to an anonymous referee for pressing me on this challenge, and to another referee for the phenomenal-to-phenomenal grounding proposal.

this phenomenal-to-phenomenal grounding response is successful depends on whether phenomenal unity is contingent. I think this relation is contingent. It seems possible that there be brain states just like mine, giving rise to discrete phenomenal experiences just like those I experience (phenomenal redness, tweeting, etc.)—but where these experiences are *not* unified.¹² Insofar as this is possible, grounding is the wrong relation to account for phenomenal unity.

If you agree about the contingency of phenomenal unity, a more promising response would take the (contingent) psychophysical bridging laws to themselves do the work of unification. Insofar as my current brain processes could have given rise to discrete conscious states, this is because there could have been bridging laws that apply to discrete regions of the brain and generate phenomenology piecemeal. Insofar as our experiences are phenomenally unified, this is because the bridging laws function in a holistic way. On a holistic account of bridging laws, there are not three separate bridging laws that attach to localized bits of brain activity, and generate (a) phenomenal reddishness, (b) phenomenal bird-shaped-ness, and (c) tweeting, which then need to be unified into the experience of a cardinal. Rather, the brain processes information in an distributed fashion, and the bridging laws attach to the workings of the brain as a whole, generating a unified cardinal-ish experience. This is itself an interesting result, and something that a thorough defense of epiphenomenalism would need to flesh out in greater detail. But the possibility of holistic bridging laws suffices to explain how discrete brain functioning could generate phenomenal unity, in the absence of phenomenal-to-phenomenal causation.

But in a sense, the details of why a quale belongs to one subject rather than to another don't need to be settled in order to distinguish qualia from KC. What's important is simply that there *is* an essential connection between qualia and subjects. There must be an explanation for why it is that qualia (or KC) are the sorts of things that can comprise subjects *at all*. We have such an answer for qualia: Qualia are essentially *experienced*, and for something to be experienced there must be an *experiencer*. By contrast, there is no reason to think that KC are eligible to comprise subjects.

One might insist that the parallel can be reintroduced by positing a further entity: the fusion of my co-conscious experiences *together with the associated Kirk's Currents*. This entity would be partially comprised by KC. So we might wonder, wouldn't this entity likewise have epistemic access to the Kirk's Currents? If so, this would reintroduce a parallel between KC and epiphenomenal qualia.

Fortunately for our argument, the fusion of KC and my mind would no more have epistemic access to KC than a fusion of my mind and the Eiffel tower would have epistemic access to the Eiffel tower. Why? First, it's not the case that every object has epistemic access to its parts. Epistemic access is something which is only had by *subjects*. And a subject plus Kirk's Currents is no more a *subject* than is a subject plus the Eiffel tower. Further, subjects only have epistemic access to those things that are presented to them in consciousness. As we'll see in §5.2, KC (unlike e-qualia) aren't suited to play the constitutive role required to afford epistemic access.

¹² These could involve disunified aspects of a single subject, or—if unity is essential to subjecthood—this could be a case where a single brain generates multiple subjects (one associated with each phenomenal “atom”).

5.2 Mental states and qualia

We can now raise the same question about the relationship between mental states and e-qualia: Why is it that e-qualia can serve as a constituent of our beliefs (and other mental states), whereas KC can't? As noted earlier, not just any old thing can serve as a constituent of my beliefs; I have to stand in the right epistemic relation to the thing. But why think that I stand in the right epistemic relation to e-qualia, but not to KC?

An item is eligible to serve as a constituent of a subject's belief only if it—like the belief—is a (mental) part of the subject.¹³ A subject's qualia are an element of her mind; a subject's KC are not. It is this shared constitutive relationship that beliefs and qualia stand in to subjects that accounts for their distinctive epistemic relationship.

We can also approach the distinctive relationship between qualia and belief from the other side, asking why beliefs are eligible to contain qualia. Note two plausible observations about beliefs (and other mental states): (a) Mental states are all essentially states *of a subject*. There can't be a belief without there being someone who believes, or a noticing without there being someone who notices. (b) Beliefs are essentially representational. In representing, beliefs fundamentally represent *to a subject*. They must *present* the content of the belief, noticing, etc. to the subject. Now, it is not the case that just any old feature (or any old epiphenomenal feature) is suited to present the world to subjects. In particular, nothing about Kirk's Currents should lead us to think that they are, by their natures, suited to present the world to subjects. Qualia, by contrast, are. Further, there is something that it's like to have contents of a belief presented to you. The world is essentially presented to me—the conscious subject—using phenomenology. Put another way: Given the way that the epiphenomenalist dualist should understand belief, beliefs are essentially phenomenal. But they are not essentially Kirk's Currents-full.¹⁴

In short, despite both being caused by and isomorphic to brain processes, the relationship between subjects and their e-qualia/KC differs: (i) KC stand in a different relationship to subjects (given the correct dualistic understanding of subjects as distinct from their brains), since KC are not constituents of the subject and the subject's mental states.¹⁵ This relationship differs because (ii) the intrinsic properties of qualia

¹³ We can clearly have beliefs with non-mental objects—in the sense of having beliefs *about* non-mental objects. But in these cases, the objects are not literal constituents of the beliefs in the sense discussed in Sect. 3. When I have a belief about the sun, the sun is not literally a constituent of my mind (setting aside naïve idealism (Yetter-Chappell 2017b, forthcoming, ms)). By contrast, when I think about my feeling of vertigo, the vertigo is not only the object of my belief, but a constituent thereof: a literal part of my mind.

¹⁴ Kirk (2005) considers a case of a zombie with KC. At some point, the laws of this zombie's world change so that e-qualia are generated instead of KC. Kirk writes, “[a]ccording to the e-qualia story the result ought to be that he becomes epistemically intimate with them” (46). In a sense this is right, but it is misleading. This suggests a picture on which there is a subject all along. Initially he lacks epistemic contact with qualia. Then we add qualia into the picture, and we further must add something to hook him up to the qualia. Contrary to this, a dualist will insist that the subject only *comes into existence* when the laws change. With that subject come the qualia that comprise both him and his beliefs, and which he is (thereby) epistemically intimate with.

¹⁵ Even if we embrace a view on which subjects are partially constituted by non-physical qualia and partially constituted by the physical brains that generate these experiences (fn 3/7), there's reason to reject KC as constituents of subjects and their mental states: while KC may be physical, they are not a part of the subvening base of our qualia (even if they are isomorphic to the subvening base). This itself gives a basis

and KC are different, such that qualia are uniquely suited to being parts of subjects and their mental states, in a way that KC are not.

6 Knowledge without causation

I have shown that epiphenomenalists can coherently make sense of the relation between qualia and our minds. But a puzzle still remains.

Consider three representative beliefs about qualia, which we'll stipulate to be true: (a) epiphenomenalist dualism is true, (b) I am feeling cold, (c) pleasure feels good. For each such belief, we can ask two further sets of questions: First, what is the causal explanation for my coming to have these beliefs? Why does my brain generate (these) thoughts about qualia? Second, can I really know the truth of these beliefs if there's no causal connection between the thing believed and my belief in it?

Since epiphenomenalists take the physical world to be causally closed, for any question of the form, "Why do our brains do P?", the epiphenomenalist's answer must be the same as the physicalist's. This is helpful in answering the first set of questions, as physicalistic opponents must grant that there are physical explanations for why I believe (a)–(c). Whatever account of these the physicalist gives, the epiphenomenalist can simply agree. Thus, the first set of questions does not pose a distinctive challenge to the epiphenomenalist.

But this exacerbates the second challenge. If there's no causal relation between the truth of dualism and my belief in it, doesn't that undermine the possibility of knowledge? Unlike the challenges discussed earlier, I do not think that this rests on a misunderstanding of epiphenomenalism. Given epiphenomenalism, the feeling of pleasure is not what *causes* me to believe that pleasure is good; the truth of dualism is not what *causes* me to be a dualist.¹⁶

This should not worry the epiphenomenalist. As Chalmers (1996) suggests, a causal theory of knowledge is simply not attractive for phenomenal knowledge. It's not just that we *can* reject the causal theory of knowledge in this domain: doing so is well motivated. It's natural for dualists to agree with Goff that phenomenal concepts transparently reveal the natures of their referents to us. (If Goff (2017) is right, the transparency of phenomenal concepts is an implicit assumption in anti-physicalist arguments.) But causal relations never allow us to transparently grasp intrinsic natures. Causal relations can reveal how things interact. But this merely gives us knowledge of structure and functional relations. To understand an intrinsic nature, we need more

Footnote 15 continued

for rejecting KC as constituents of subjects (even for those inclined to treat neural states as constituents of subjects). But more importantly, even if KC were parts of subjects, *they are not part of the m-part (the conscious part) of subject*. And it is the m-part which is relevant to epistemic contact on the dualist picture I have constructed.

¹⁶ Had my brain been structured differently, this might have led me to a (false) belief in physicalism. Yetter-Chappell (2017a) sketches such a neural architecture. Yetter-Chappell (2019) argues that—while *physicalists* should take such agents to be more rational than us—dualists should take agents with such a cognitive structure to be systematically irrational.

than how it affects other things. We need the thing itself to be manifest within our minds.¹⁷

While we would ultimately like a complete account of (i) when the causal theory of knowledge is appropriate and (ii) what the alternative theory of knowledge is, the case I've presented is sufficient to cast doubt on the applicability of the causal theory in the phenomenal domain. And this in turn is enough to undermine the charge that the epiphenomenalists' position on knowledge is *paradoxical*.

But even if it is not paradoxical, it may still seem bizarre and unsettling to think that if only the bridging laws had been different, I would have believed (c') pain is good, instead of (c) pleasure is good. How could my phenomenal beliefs be a matter of such chance?

Let's start by considering the chanciness of belief (b). My cold-fibers are firing,¹⁸ causing me to have a cold experience. Had the bridging laws been different, cold-fibers firing would have caused an experience of heat. Given the constitutional account of phenomenal concepts endorsed in §3, phenomenal experiences can be constituents of our concepts and the judgments that contain these concepts. When I think 'I'm feeling cold,' the very experience I'm referring to is a part of my thought. The envisioned change to the bridging laws does not change the structure of my thought, but it does change the content of what I attribute to myself. So when I'm feeling cold, it is a cold experience that my thought attributes to me. If the bridging laws instead make me feel hot, they correspondingly make it a hot experience that I'm self-attributing.¹⁹

The puzzling cases for the epiphenomenalist are ones in which my brain generates thoughts that do more than self-attribute occurrent experiences: where a further claim is being made of the experience, which could fail to co-vary with tweaks to the bridging laws. Belief (c) provides such a case.

It might seem strange to think that I could easily have believed that pain is good (had only the bridging laws been different). But this is the flip side of something that is a clear benefit of epiphenomenalism. Intuitively, there is no contradiction in two people being physically alike, but color-inverted. Physicalists cannot account for this intuitive datum, and this is often taken to be a challenge to physicalism. But this is simply another instance of the same point that is now being taken to be a counterintuitive consequence of epiphenomenalism: the contingency of the connection between neural states and experiences.

¹⁷ The epiphenomenalist view of transparency is nuanced: While we can agree with Goff that in employing a phenomenal concept, the nature of the concept is (in some sense) *available* for us to grasp—by virtue of being a constituent of our minds—employing these concepts is clearly not sufficient to *actually grasp* the nature. Consider the case where (due to unfortunate bridging laws) I think "pain is good". *Pain* is laid bare to me, in my mind. The *badness* of pain is laid bare to me. And yet—because of the physical structure that generates thoughts—I do not attend to this badness, instead formulating a sort of word-salad thought putting together the (bad) pain which I experience with the concept of goodness. By contrast, in the good case, the badness of pain is not only laid bare to my mind, I can attend to it and reflect on its badness. (I'm taking for granted here that our concepts of good and bad are not constitutively linked to pain/pleasure sensations).

¹⁸ Stipulate that this is the neural correlates of cold experiences, whatever that might be.

¹⁹ Chalmers (2003) offers a fleshed-out theory of phenomenal concepts which renders this clear.

We cannot have it both ways.²⁰ We cannot reasonably insist that experiences are *sometimes* contingently connected to physical states, and other times are not. The physicalist might enthusiastically take this to provide a basis for rejecting the conceivability of inverted spectrum cases: “Sure, it might have seemed like inverted spectrum cases were conceivable—but then we found out what else that would entail.” (This is the style of reasoning Kirk (2005, 2008) uses in arguing from Kirk’s Currents to the “inconceivability of zombies”.)

The epiphenomenalist could simply flip this reasoning: granting that the possibility of (c’) is counterintuitive, but insisting that it’s more important to hold onto the conceivability of inverted spectrum cases. But I think the epiphenomenalist has a better move available. Epiphenomenalism does a better job of respecting our intuitions *even in cases like (c)*.

The idea that we *might actually* be wrong about our phenomenal judgments is intuitively preposterous. But the epiphenomenalist does not think this. Suppose I reflect on the pain in my arm and think of it “this feels bad”. Pain—with its horrible qualitative nature—is a part of my mind. The badness of pain is plausibly not only a part of my mind but a part of my evidence base. (Albeit one that I can only access provided my brain functions so as to cause such a thought.) But when I *do* reflect and judge that pain is bad, the badness of pain is before me, there as part of my mind. So it would be wrong to think that I might *actually be* wrong in making this judgment. Rather, the epiphenomenalist thinks that we *could have been* wrong in our phenomenal judgments. I could have lived in a world with unfortunate bridging laws. In this world, I would be a conscious agent who thinks in word-salad: asserting that *manifestly* wonderful experiences are bad. While this may be unsettling, it also seems intuitively *right*. There is nothing inconceivable about people who *behave* just like us in responding to tissue damage (including at the neural level), and yet feel pleasure when it occurs—just as there is nothing inconceivable about having red experiences when looking at a green leaf. Epiphenomenalism gets logical space right; those who would deny the contingency of our phenomenal judgments get it wrong.

It would be nice if there were some basis for holding that while world c’ is *possible*, it’s not *likely*. If there were a basis for this, we could get logical space right, but still hold that it’s not a *mere* coincidence that we don’t inhabit such a world. One basis for doing this would be a belief in a rational and benevolent God. (I’m inclined to think a fine-tuning argument based on fine-tuning the bridging laws is the strongest argument in favor of God.) Alternatively, one might think that certain possibilities (e.g. those with bridging laws that facilitate true beliefs) are simply a priori more likely.²¹ Either move would allow the epiphenomenalist to hold onto the intuitive sense that—although we could have been in the bad case—it’s not *mere* chance that

²⁰ One might think that interactionist dualism would allow us to hold that—though experiences are contingently connected to physical states—they cannot come apart in ways that generate irrational thoughts like “pleasure is bad”. But this is just as much a problem for interactionists as it is for epiphenomenalists, for there’s no reason to think that it would be *impossible* for pleasure to cause the thought “this is horrible”. Such a thought may be irrational, but it is not *inconceivable* that pleasure *cause* such a thought. See Wright (2015) for a defense of this point.

²¹ Chappell (2017) proposes an analogous move in responding to causal inefficacy objections to non-naturalist moral realism.

we've avoided this. Alternatively, we may simply regard ourselves as lucky to live in a world with sensible bridging laws, just as we may regard physical fine-tuning as fortuitous.

In sum: (1) Epiphenomenalists face no problem explaining why my brain generates thoughts about qualia (either in general or particular ones): they can appeal to the very same explanations as physicalists. (2) The lack of causal connection between qualia and my beliefs about qualia does not undermine my ability to have knowledge of qualia, since a causal theory of knowledge is not suitable for qualia. Rather, I directly grasp the nature of my qualia in experiencing them. (3) Epiphenomenalists are not committed to holding that we might *actually be* wrong about qualia, only that we *could have been* wrong about them. And (4) this intuitively gets logical space right.

7 The remaining paradoxes

In Chalmers's (1996) presentation of the paradox of phenomenal judgment, he considers three related paradoxes: arguments that epiphenomenalists cannot account for (i) knowledge of qualia, (ii) memories of qualia, and (iii) reference to qualia. I've shown that the epiphenomenalist can account for phenomenal knowledge. What of the remaining paradoxes?

In Sect. 7.1, I'll consider the challenge to memory, defending Chalmers's (1996) response. In Sect. 7.2, I'll consider the challenge of reference, distinguishing three related challenges to phenomenal reference. The earlier discussion of the nature of subjects does not address these challenges, and one of the three challenges is made more acute by embracing a phenomenal conception of subjects. Nevertheless, I'll argue that these challenges can be met, and none are paradoxical.

7.1 Backward-looking mental states

I've argued that the most natural way for dualists to understand the relationship between sensations, mental states, and subjects is in a nested way, such that your mental states and your sensations are all parts of *you*, present within your mind. In cases where we have a mental state about an experience, the sensation is a part of the mental state, which is in turn a part of the subject's mind. Mind and all of its constituents, on this view, are phenomenal. So although there may be "distance" between mental states and your brain, there is no "distance" between your (conscious) mind and your (conscious) mental states. We don't need causation to span the gap, because there isn't a gap.

This case is most naturally made for mental states whose natures can be entirely understood by looking at the present moment: beliefs, desires, wonderings, etc. But backward-looking mental states—rememberings, phenomenal comparisons across time—might seem difficult to fit within this phenomenal model of mental states. On standard models of epiphenomenalism, past phenomenal states cannot affect current mental states. But contact with the past is precisely what backward-looking mental states require. As Kirk (2008, p. 76) writes:

Epistemic contact...involves cognitive processes such as conceptualization and the storing and retrieving of information, which in turn involve the causation of changes and the persistence of unconscious items. ... If information is stored about an event, that must leave traces which can have effects later, when the subject is recalling or otherwise using the stored information; but these traces are not normally conscious. ... Epistemic contact also involves causation and unconscious persistence, because it depends on conceptualization, which requires more or less persisting cognitive structures contributing causally to the subject's ability to group things together.

So (how) can the phenomenal model of mental states make sense of backward-looking mental states?

Memory essentially relies on information about the past being *currently* accessible to my mind. My current temporal parts only have epistemic contact with the past via current epistemic contact with this stored information. So there are two things that are essential for such backward-looking mental states: (1) information about the past must be stored, and (2) the subject must be put in epistemic contact with this information. Insofar as one embraces the (dualist-friendly) phenomenal nesting view of epistemic contact with experiences, this is what's required for (2). But to account for (1), the epiphenomenalist must appeal to the physical storage of information in the subject's brain—since, by hypothesis, past phenomenal experiences are causally inefficacious, and so incapable of storage. For some purposes, (1) might be all that we're interested in (e.g. for understanding how learning works, or what's distinctive to memory). But (1) only gives us *conscious access* to our (past) experiences insofar as (2) is met.

Here's how it works for the epiphenomenalist: Traces of the neural correlates of the original experience are laid down in the brain. These are subsequently used to generate the conscious experience directly involved in the memory—an experience qualitatively akin to the original experience. Our epistemic contact with past phenomenology is the result of this subsequent epistemic contact with our current phenomenology. And this occurs in precisely the same manner as with noticings: the subsequent phenomenology is a *constituent* of my current mental state of remembering. Thus, for an epiphenomenalist, genuine memory involves (a) an experience E_1 occurring at t_1 , (b) information I about the neural correlates of E_1 being recorded in the brain, (c) information I subsequently being used at t_2 to generate a new experience E_2 , which is qualitatively similar to E_1 , (d) E_2 being a constituent of an appropriate mental state of the subject.²² As on a traditional conception of memory, a causal connection between the past and the present is essential. But while a traditional conception takes a causal connection between E_1 and E_2 to be essential, the epiphenomenalist holds that the causal connection is between I and the experience E_2 . Merely apparent memories—as Parfit's quasi-memories of being Napoleon (1984)—involve an experience qualitatively like E_2 , but lack (a), (b), and/or (c). This epiphenomenalist conception of memory is a

²² Broad and narrow views of mental states (ftn 7) will differ in how they make sense of memory. On a narrow view, initial experience E_1 and E_2 —presented to consciousness as an apparent memory—might be mental states, while information I and the causal process that leads to the formation of E_2 would merely be the causal underpinnings of the mental state. Apparent memories are mental states, while the information storage and causal processes that render apparent memories genuine are a non-mental addition. Whereas on a broad view, information I is a part of the mental state.

familiar one, and has been defended by Chalmers (1996, pp. 200–201) and Robinson (2016, p. 55).

Epiphenomenalists' account of memory must diverge from the standard model. From the epiphenomenalist's perspective, this account gives us all that we have reason to posit, and preserves a basis for distinguishing genuine memories from merely apparent ones. I don't think that all readers will find the epiphenomenalist conception of memory compelling—just as libertarians about free will don't find compatibilism adequate. Rudd (2000, p. 60) complains that “[l]ike many other theories that claim to ‘save the appearances’ ... we are left with the appearances, but the reality of memory has been lost.” But even granting that there are grounds that non-epiphenomenalists might have to dispute the epiphenomenalist's account of memory, the epiphenomenalist-friendly accounts are sufficient to dissolve the *paradox*. For there to be a paradox, there must be a challenge that shows that even by the lights of the theory itself, there is something that undercuts the theory. And the objection from memory does not do this.

7.2 Talking about qualia

Beyond internally making phenomenal judgments and having phenomenal memories, dualists of all stripes spend a lot of time saying and writing things like “Tell me everything physical there is to tell about what is going on in a living brain, ... [and] you won't have told me about the hurtfulness of pains, the itchiness of itches, pangs of jealousy...”²³ If epiphenomenalist dualists were unable to account for such public references being meaningful, they would be unable to publicly assert the truth of epiphenomenalism. Foster (1991, p. 191) put the problem forcefully:

I cannot even see how, on the epiphenomenalistic view, our language ... makes semantic contact with the mind at all. ... It is true, of course, that each language-user may mentally interpret each term as signifying a certain kind of (dualistically conceived) mental item. But I cannot see how such private interpretations could have any bearing on the objective meaning of the terms, as employed in speech and writing, if, with respect to this employment, they are causally idle.

I think that there are actually three challenges that might be raised on these grounds. An intuitive picture of reference has it that there is a causal connection between (i) the object of reference and the referring subject, (ii) the utterance and the object of reference, and (iii) the referring subject and the utterance. One might object to the epiphenomenalist's picture of reference on grounds that any of these three natural assumptions about reference breaks down.

First, one might object that reference requires a causal connection between the object of reference and the referring subject doing the referring. Since there is no such causal connection between qualia and referring subject, there cannot be reference to qualia. This variant of the challenge can be responded to in precisely the same manner as the original paradox: the connection between subject and referent needn't

²³ A representative quote from Jackson's “Epiphenomenal Qualia” (1982).

be causal, it can be constitutive. And there is a constitutive relationship between the (phenomenal) subject and qualia.

A second objection would hold that there must be a causal connection between the utterance (or the producer of the utterance) and the referent: the referent must be at the end of a causal chain leading to the utterance. Since the utterance is a physical result of my vocal cords vibrating (and the producer of the utterance is my body), there is, for the epiphenomenalist, no causal connection leading from qualia to my utterance “there’s something it’s like to feel pain”.

As I understand it, this is at least part of the challenge that Foster raises, and is also the challenge that Chalmers (1996) considers. This problem is not solved by embracing a dualistic conception of subjects. But, as Chalmers (1996) argues, this challenge can be responded to by rejecting the causal theory of reference.

There seems no principled reason why reference to an entity *requires* a causal connection to that entity... In referring to an entity, all that is required is that our concepts have *intensions* (in particular, primary intensions) that the entity might satisfy. ... It happens that for many of our concepts, primary intensions are categorized causally... But there is no reason why a primary intension *has* to work this way. ... Even a brain in a vat might have concepts with primary intensions, despite its causal isolation... Again, the *constitution* of a primary intension is independent of such causal connections. (201–202)

We can refer to mathematical entities despite not being causally related to them, so a causal relation between referents and utterances/utterers is not always essential for reference. The epiphenomenalist takes it that qualia are another such case. But there’s another challenge lurking in this vicinity that Chalmers’s response doesn’t address—which also seems to be contained within Foster’s argument.

This objection holds that there must be a causal connection between the conscious thought the subject intends to convey and the utterance. The idea would be that—for my utterance “this hurts” to be meaningful—there must be a causal connection between (i) the thought that I consciously grasp and want to convey and (ii) the utterance leaving my mouth. This challenge is made *more* acute by embracing a phenomenal conception of subjects and their thoughts—for it doesn’t just pose a challenge to epiphenomenalist’s ability to make sense of reference to phenomenology, but to all reference. For the epiphenomenalist, my utterance “the tea is ready” is no more the causal result of some thought I intend to convey than is my utterance about pain.

As in the cases of memory and knowledge, epiphenomenalists must embrace a different conception of what’s required for reference than the picture we started with. Start by considering what’s intuitively *essential* for successful communication. Utterances successfully convey semantic content when they result in hearers understanding what the speaker intends to convey. Communicative exchanges are successful when (i) the hearer comes to understand what the speaker intended to convey, *and* (ii) the speaker can reliably count on the speaker coming to successfully understand. Both these conditions are satisfied for epiphenomenalists. The utterance affects the hearer’s brain, causing them to have a conscious thought “the tea is ready” or “that hurts”. Because these utterances were produced by the same brain that also produces the speaker’s conscious thoughts, both speaker and hearer come to understand the same thing. And this

is something that the speaker can—even as a committed epiphenomenalist—expect and count on.

Still, one might wonder how utterances come to be meaningful. If there is no causal connection between utterances and the meanings I intend to convey, it might seem that the two don't "stick together" in the right way. (Cf. Rudd, 2000). Clearly, utterances (and writing) for the epiphenomenalist do not have original intentionality. (This seems clearly right to me independently of epiphenomenalism.) Since the conscious subjects to whom we attribute utterances are not causally related to the utterances themselves, we might wonder how the latter derive their intentional content.

This question does not seem to me a deep question, but rather a matter of convention in how we describe things. There are deep questions of what the speaker means to convey, what the listener interprets them as meaning, and whether these two things match up. (We've already established that epiphenomenalists have no trouble accounting for this.) But utterances are only meaningful in that we take them to be meaningful. Such attributions of intentionality are useful. But there's no reason to think that there are deep facts about derivative intentionality that outstrip the conventions we use in attributing such meaning. Among the conventions we might take to confer derived intentionality: (i) The utterance is interpreted as meaning something by listeners, (ii) the utterance would be reliably interpreted as meaning something by all (or a sufficient number of) members of the linguistic community and (iii) the same physical state which produces the utterance also produces a conscious intentional state (e.g. the same brain state causes "that hurts" to be uttered and a conscious thought "that hurts" to be had by a conscious subject).²⁴

As with the epiphenomenalist's account of memory, it's not essential (to the task of addressing the paradox) that non-epiphenomenalists find the account of reference compelling. What matters is that it's compelling—internally consistent and well-motivated—for the epiphenomenalist. With such an account in hand, we can see that the paradox of phenomenal reference isn't so paradoxical after all.

8 Conclusion

We've looked at a range of putative paradoxes for the epiphenomenalist, stemming from the alleged inability of epiphenomenalists to account for epistemic contact with our phenomenology: paradoxes of self-knowledge, memory, and reference. We've seen that none of these is truly paradoxical. When it comes to our occurrent phenomenal judgments, the appearance of paradox only arises when one combines the epiphenomenalist's picture of qualia with a physicalistic conception of subjects and their introspective knowledge. Kirk (2005, 2008) highlights the challenges that must be met to respond to the paradox, and why simply appealing to acquaintance is not sufficient as a response. But the fundamental flaw remains the same: not the theory of consciousness, but the mismatch between the theory of consciousness and the account of subjects and their cognitive lives. We should not be surprised that trying to squeeze

²⁴ Option (iii) strikes me as most plausible given an account of subjects as partially constituted by non-physical qualia and partially by physical brains.

a dualistic picture of qualia into a physicalistic conception of the rest of the mind will have bizarre results. The lesson to take from this is not that epiphenomenalism is false, but that epiphenomenalist dualists should be “thoroughgoing dualists”: embracing the importance of the non-physical not only in accounting for qualia, but also in accounting for subjects and their cognitive lives.

While the paradoxes of memory and reference show that epiphenomenalist accounts of memory and reference must diverge from the standard models, the epiphenomenalist has accounts of each that—by their own lights—are sufficient to explain the phenomena. (And these accounts don’t merely involve thumping fists against tables and insisting that the phenomena don’t need to be accounted for. They are accounts that genuinely capture—from an epiphenomenalist’s perspective—all that we truly have reason to posit.) And *that* is sufficient to dispel the paradox.

Thus we see that what is widely regarded as the most devastating objection to epiphenomenalism is no problem after all. There is no paradox of phenomenal judgment.

Acknowledgements Thanks to Richard Yetter Chappell, Robert Kirk, and two extremely helpful referees for their comments on earlier drafts of this paper. Thanks also to participants of the New Directions Non-physicalist Theories of Consciousness Conference and University of York Mind & Reason research group for much fruitful discussion.

References

- Balog, K. (2012). Acquaintance and the mind-body problem. In C. S. Hill & S. Gozzano (Eds.), *The mental, the physical*. Cambridge University Press.
- Chalmers, D. (1996). *The conscious mind*. Oxford University Press.
- Chalmers, D. (2003). The content and epistemology of phenomenal belief. In Q. Smith & A. Jovic (Eds.), *Consciousness: New philosophical perspectives*. Oxford University Press.
- Chappell, R. Y. (2017). Knowing what matters. In P. Singer (Ed.), *Does anything really matter? Essays on parfit on objectivity*. Oxford University Press.
- Chappell, R. Y. (2019). Why care about non-natural reasons? *American Philosophical Quarterly*, 56(2), 125–134.
- Descartes, R. (1641/1996). *Meditations on the first philosophy* (J. Cottingham, Trans.). Cambridge University Press.
- Foster, J. (1991). *The immaterial self: A defence of the Cartesian dualist conception of the mind*. Routledge Press.
- Frege, G. (1918/56). The thought: A logical inquiry. *Mind* (65) (P. T. Geach, Trans.).
- Gertler, B. (2011). *Self-knowledge*. Routledge Press.
- Gertler, B. (2012). Renewed acquaintance. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness*. Oxford University Press.
- Goff, P. (2017). *Consciousness and fundamental reality*. Oxford University Press.
- Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32(127), 127–136.
- Kirk, R. (2005). *Zombies and consciousness*. Oxford University Press.
- Kirk, R. (2008). The inconceivability of zombies. *Philosophical Studies*, 139, 73–89.
- Papineau, D. (2002). *Thinking about consciousness*. Oxford University Press.
- Parfit, D. (1984). *Reasons and persons*. Oxford University Press.
- Pauen, M. (2006). Feeling causes. *Journal of Consciousness Studies*, 13, 129–152.
- Pitt, D. (2011). Introspection, phenomenality, and the availability of intentional content. In T. Bayne & M. Montague (Eds.), *Cognitive phenomenology*. Oxford University Press.
- Robinson, W. (2006). Knowing epiphenomena. *Journal of Consciousness Studies*, 13, 85–100.
- Robinson, W. (2012). Phenomenal realist physicalism implies coherency of epiphenomenalist meaning. *Journal of Consciousness Studies*, 19, 145–163.

- Robinson, H. (2016). *From the knowledge argument to mental substance: Resurrecting the mind*. Cambridge University Press.
- Rudd, A. (2000). Phenomenal judgment and mental causation. *Journal of Consciousness Studies*, 7(6), 53–66.
- Shoemaker, S. (1975). Functionalism and Qualia. *Philosophical Studies*, 27, 279–315.
- Strawson, G. (2003). What is the relation between an experience, the subject of the experience, and the content of the experience? *Philosophical Issues*, 13, 279–315.
- Strawson, G. (2017). *The subject of experience*. Oxford University Press.
- Wright, J. (2015). On James's argument against epiphenomenalism. *William James Studies*, 11(1), 69–85.
- Zimmerman, D. (2011). From experience to experiencer. In M. C. Baker & S. Goetz (Eds.), *The soul hypothesis: Investigations into the existence of the soul*. Continuum Press.
- Yetter-Chappell, H. (2017a). Dissolving type-B physicalism. *Philosophical Perspectives*, 31(1), 469–498.
- Yetter-Chappell, H. (2017b). Idealism without god. In T. Goldschmidt & K. L. Pearce (Eds.), *Idealism: New essays in metaphysics*. Oxford University Press.
- Yetter-Chappell, H. (2019). Idealization and problem intuitions: Why no possible agent is indisputably ideal. *Journal of Consciousness Studies*, 26, 270–279.
- Yetter-Chappell, H. (ms). *The View from Everywhere: Realist Idealism Without God*.
- Yetter-Chappell, H. (forthcoming). Get Acquainted With Naïve Idealism. In R. French & B. Brogaard (Eds.), *The Roles of Representations in Visual Perception*. Synthese Book Series.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.