

Hidden Qualia

Derek Shiller

(Penultimate Version)

1 Introduction

The doctrine of mental infallibilism, which holds that we cannot be mistaken about our own mental states, has had a bad century. Its popularity has declined as we have come to understand more about the complex subliminal processes that shape our cognitive lives. Despite this, the doctrine remains compelling when restricted to a limited domain: it is deeply intuitive that we are *nearly* infallible when it comes to the nature of our own current conscious experiences. I will argue, however, that we should be open to the possibility that there is more to our current conscious experiences than we are aware of.

Before I can be more specific about my proposal, I will need to lay out two fairly intuitive assumptions. These assumptions are controversial¹ and are subject to reasonable doubts. Nevertheless, they are popular and it is worth exploring their consequences.

First, I will assume that there is an important sense in which a mental state is *conscious* just in case it has a *phenomenal character*. Phenomenal characters are the qualities of mental states that collectively constitute what it feels like to be us. The experience of appreciating the redness of a ripe tomato, the painfulness of stubbing one's toe, the satisfaction of accomplishing one's goals, and the sense of déjà vu each possess their own distinct phenomenal characters.

I will use the familiar term 'qualia' for the phenomenal characters of conscious states. This term has a number of rich associations in the philosophical literature. According to Dennett's (1988) summary, qualia are (or would need to be) ineffable, intrinsic, private, and directly apprehensible

¹The first assumption is rejected by those physicalists who deny that there is any deep sense in which our experiences have phenomenal characters. They might hold, for instance, that we attribute phenomenal characters to our mental states as part of an over-simplified model of cognition (Graziano, 2013), and that there is nothing in reality that fits the model (Dennett, 1988). The second assumption is rejected by those physicalists who deny that the phenomenal characters of our mental states are independent of our beliefs about them, as in the higher-order thought theories developed by Rosenthal (1990) and Carruthers (2003).

mental qualities. While Dennett captured something important about how we ordinarily conceive of qualia, I do not mean to build it into the concept of qualia that they must fit his description. Instead, I intend for the content of the concept to be fixed by its prototypical examples. Qualia are whatever kind of mental qualities we associate with experiences of redness, pain, satisfaction, and déjà vu. This approach leaves the veracity of our intuitive assumptions open to investigation.

Second, I will assume that qualia are in no way either constituted by or metaphysically dependent upon our beliefs about our own experiences. What it is to feel a certain way – for a state to have a particular phenomenal character – is not even partly to believe that one feels a certain way. Beliefs about our experiences are cognitive in a fashion that the experiences themselves are not. There may still be necessary connections between our qualia and our beliefs about our experiences, but such connections are not inherent in the nature of qualia itself.

My proposal is that we are not justified in believing that we lack what I will call, following David Papineau (2002), ‘hidden qualia’. Hidden qualia are qualia that are such that we cannot tell that we have them simply by introspection. More precisely, I mean that qualia are hidden if we are not disposed to form or alter our beliefs about them as a result of introspection.² Any qualia that we only know about from the careful study of brain scans would be hidden. I will say that qualia are ‘introspectible’ if we can tell that we are having them by introspection.

Since hidden qualia are phenomenal characters, they make a contribution to what it is like to be us.³ Hidden qualia need not be any less vivid or substantial than any other qualia. Hidden qualia are only hidden from our faculties of belief. Since we normally assume that we know by introspection more or less what it feels like to be us, the possibility of hidden qualia suggests that we might have radically mistaken beliefs about the character of our own conscious experiences.

My argument that we are not justified in believing that we lack hidden qualia will involve three steps, which I will defend in the three main sections of this paper. First, I will present four arguments for thinking that hidden qualia are possible. Second, I will present two arguments for thinking that we actually have hidden qualia. The inconclusiveness of these arguments reflects the modesty of my aims. I do not hope to establish that we have hidden qualia, only that we should

²No causal relation is necessary: a pre-established harmony between our introspective faculties and our qualia would be sufficient for those qualia not to be hidden as long as we would still be poised to form beliefs about them as a result of introspection. For this reason, epiphenomenal qualia are not automatically hidden.

³Carruthers (1989) talks about ‘unconscious experiences’, which are similar to what I am calling ‘hidden qualia’, but he claims they make no contribution to what it is like to be us. For that reason, hidden qualia need not be involved in unconscious experiences.

not rule out the possibility. I think the best prospects for justifying belief in the non-existence of hidden qualia come from recent work on the epistemology of epiphenomenal qualia. So third, I will object to several strategies that might be borrowed from that literature for justifying belief in the non-existence of hidden qualia.

2 The Possibility of Hidden Qualia

Many philosophers have taken it for granted that all of our qualia are introspectible. Our paradigm examples of qualia, and those through which we acquired the notion, are all introspectible. They must be if we are able to attend to and remember them. This means that there is an important epistemic tie between qualia and introspection, and perhaps this tie plays some role in explaining the pull toward thinking that qualia must be introspectible. We should be wary, however, about inferring that all qualia must be introspectible from the fact that we are only able to introspect introspectible qualia. We can grant that qualia are the sorts of things we learn about through introspection without concluding that we must be able to introspect all of them.

Even though I am assuming that qualia are not to be analyzed even partly into higher-order beliefs, it is still possible that hidden qualia are impossible because qualia must be introspectible as a result of some sort of metaphysical law. In this section, I will offer four arguments that hidden qualia are at least possible. If any one of these arguments succeeds, then introspectibility is not a necessary feature of qualia.

The first argument relies on the premise that *partially* hidden qualia are possible. Hidden qualia are qualia that we cannot tell we're having. Partially hidden qualia, on the other hand, are qualia that we can tell we're having only with some difficulty. Unlike hidden qualia, partially hidden qualia are *poised* to influence our higher-order beliefs – it just takes effort or the right context for them to do so.

We might have partially hidden qualia when our attention is directed away from our experiences. If attention is required for our experiences to influence our higher-order beliefs, then qualia that escaped our attention would be partially hidden qualia. For example, the sound of the air conditioning in the background may have no impact on your beliefs about your experiences when you are trying to pay attention to an engaging lecture. It is partially hidden because it would become introspectible if you were to attend to it. While the existence of unattended qualia

is controversial (Prinz, 2000; Papineau, 2002; Schwitzgebel, 2007), it is intuitively plausible that we have some qualia that we do not attend to.

Sometimes contextual factors make it especially difficult to direct our attention to a particular stimulus. This occurs, for instance, when a distracting stimulus draws our attention away from a more subtle one. In such cases, it is plausible that although it might be especially difficult to direct our attention to a given quale, that quale still constitutes part of what it feels like to be us. Ned Block (2007; 2008) has even suggested that the phenomenon of visual extinction may represent an extreme example of this – a neurological deficiency makes it impossible to draw one’s attention away from certain kinds of distracting stimuli.

Given the premise that partially hidden qualia are possible, we can argue as follows. The difference between partially and fully hidden qualia is a matter of degree. As the amount of effort necessary for introspecting a quale gets higher, the quale will come closer to being fully hidden. Total hiddenness is an endpoint on a scale. There is no natural prior point on that scale to start denying possibility to qualia that are *only hidden to that degree*. It would be unacceptably arbitrary for partially hidden qualia to be possible up to some point on this scale, but no further. If it is agreed that partially hidden qualia are possible, then we should expect hidden qualia to be so possible as well.

The second argument is a relative of the first. If we hold that every quale must have the capacity for some effect on our higher-order faculties, we invite the question: what sort of influence must they be capable of? Upon introspecting a rose-red quale, we might come to believe that we are experiencing a rose-red quale, a red quale, a colored quale, or some visual quale or other. Introspection might also lead us to different levels of confidence about our higher-order judgments: we might feel totally certain about our experiences or we might only have a slight inkling that we’re having a given experience. If we think that hidden qualia are impossible, we are forced to suppose that there is some minimal level of specificity of content and certainty necessary for those beliefs.

We do not explicitly form beliefs that capture every phenomenal aspect of our experiences. Beliefs fit our experiences under concepts, and our concepts are inadequately precise to capture our experiences in full (Tye, 1995; Byrne, 1997). Furthermore, when pressed we are often not very confident about many aspects of our experiences.⁴

⁴Eric Schwitzgebel (2011) provides a range of examples in which introspection produces uncertainty or pervasive disagreement about the nature of our experiences.

This argument has the same form as the last. It is implausible that there are any minimal requirements for a quale's effect on the certainty or the specificity of our higher-order beliefs. Any such minimal requirements would be arbitrary. If there are no minimal requirements, then there are no requirements at all. Hence qualia need not be introspectible.

The third argument focuses on the capacities of non-human animals.⁵ One common criticism of the higher-order theories of consciousness (such as in Block 1995) is that they imply that most non-human animals are not conscious. Many philosophers have found it plausible that most non-human animals lack the cognitive sophistication to form beliefs about their conscious states, either because they lack the right concepts, or because they do not have the right cognitive organization to recognize their own mental states. If they nevertheless do have qualia, then their qualia must be hidden, and if any animals actually have hidden qualia, then it must be possible.

Finally, there is an argument from Hume's dictum. Hume's dictum holds that (in Jessica Wilson's summation) "there are no metaphysically necessary connections between distinct intrinsically typed entities." (Wilson, 2010, 598) If qualia and beliefs are distinct intrinsically typed entities, it will turn out that any given quale can exist without any particular effect on higher-order beliefs.

Plausibly, qualia are intrinsically typed entities. It is rather less plausible that beliefs, qua beliefs, are also intrinsically typed entities; many psychosemantic theories hold that the content of a belief is a matter of its etiology or its relation to other mental states and to the world. This does not undercut the argument, for we don't need beliefs to be intrinsically typed entities in order to use Hume's dictum to show that qualia need not be introspectible. So long as beliefs and qualia both have distinct intrinsically typed parts or are subkinds of distinct intrinsically typed kinds, then beliefs must be distinct from qualia in the relevant sense. If beliefs are distinct mental states from qualia, Hume's dictum will deny them any necessary connection.⁶

If Hume's dictum is true and if beliefs and qualia really are distinct entities, then it should be metaphysically possible to have one without the other. Hence, our qualia would not necessarily need to produce any particular higher-order beliefs.

⁵Papineau (2002) gives a version of this argument in defense of the possibility of hidden qualia.

⁶Even leaving aside higher-order theorists, not everyone agrees that qualia and beliefs about qualia are distinct. One strain of the phenomenal concept strategy (developed, for instance, in Hill and McLaughlin 1999 and Chalmers 2003) holds that qualia are constituents in some of our beliefs about them. Nevertheless, we might still use Hume's dictum to demonstrate the possibility of hidden qualia so long as the relevant beliefs contain constituents apart from the quale, that are themselves intrinsically typed and distinct from the quale. It is exceedingly plausible that beliefs have a part that goes beyond whatever qualia they are about, and that Hume's dictum will tell us that at the very least, qualia could exist while whatever else would be required for a further belief does not.

3 The Actuality of Hidden Qualia

In the last section, I argued that we should think hidden qualia are possible. Possibility is cheap; the vast majority of possible things don't exist. In this section, I'll present two different arguments for believing that we probably have them. The first argument, the *exceptionality argument*, concerns the probable limitations of our faculty of introspection. The second argument, the *argument from varieties*, concerns the variety of kinds of hidden qualia that we might conceivably have.

3.1 The Exceptionality Argument

Vision was very important for our ancestors. The ability to navigate the environment by sight helped them to survive in a competitive world. While our vision is quite good, it is imperfect. There is much around us that we cannot see, including things that are very small, that do not reflect light of the right part of the spectrum, or that are blocked by other things.

There is a good reason for this imperfection. Our perceptual capacities are limited by both physical and evolutionary constraints. We should not expect our capacities to be much better than what was necessary to help our ancestors thrive. We cannot see small objects at great distances or distinguish light of ultra-violet wavelengths in part because these capacities were not selected for in our ancestors' ecological niche, and did not free-ride on anything that was.

While our faculty of introspection may work rather differently from our faculties of external perception, we should expect that it is also limited both by our ancestors' needs and by the physical constraints imposed by the structure of their brains. A brain doesn't automatically produce beliefs about its own states. This capacity surely requires sophisticated adaptations that likely come with *some* associated cost. Just as our capacities for external perception are imperfect because it was not sufficiently advantageous to our ancestors to be able to perceive everything in their environment, it is likely that our faculty of introspection is imperfect because it was not sufficiently advantageous to our ancestors for it to be infallible.⁷ If it is limited, and hence we do not have accurate beliefs about all of our qualia, it is a priori quite plausible that one of its imperfections is an inability to detect certain kinds of qualia.

⁷It is conceivable that a flawless faculty of introspection might have resulted from a neural reorganization that did not impose any metabolic cost. (Thanks to an anonymous reviewer for this suggestion.) Such a reorganization would not need to earn its keep. Nevertheless, the reorganization would not spread throughout the population unless it were specifically selected for.

There are two straightforward ways to resist this argument. First, one might make a case for thinking that there is some significant advantage to be gained in knowing about all our qualia. The argument rests on the supposition that some of our qualia might have been insufficiently important to our ancestors' well-being to exert any pressure on the development of our introspective capacities. If this were false – if all our qualia were worth knowing about – then it makes sense that we might have developed introspective capacities that were sufficiently fine-tuned to capture all of our qualia.

It is *prima facie* dubious that our ancestors would have benefited from knowledge of all of their qualia. If, as is often thought, many animals have qualia that they cannot introspect, then it must be possible for a creature to get along pretty well without knowing about some of its qualia. Humans may have more use for some knowledge of their qualia than other animals, but there is no reason to think that we should need to know about absolutely all of it. Without a clear rationale of why qualia should have to be useful to know about, we should be open to the existence of many kinds that are not. When I consider possible sources of qualia in the next argument, I will provide several examples of kinds of hidden qualia that would be useless to know about.

Second, one might reply that the same mechanisms used to introspect some of our qualia also allow us to introspect all of it. The pressures that led our ancestors to be able to introspect some of their qualia also incidentally provided an ability to introspect it all.

However, it is far from obvious that our ancestors' brains should have been organized in a way that would allow one set of pressures to have such an effect. If qualia arise in diverse ways and in different parts of the brain, then we might expect that acquiring the capacity for flawless introspection would have required a long series of specific targeted changes – each of which would need to stand on its own. Until we understand the mechanisms by which qualia are produced and introspected, we should doubt that the most straightforward neurological changes that would have would have served our ancestors' needs would have allowed them to introspect absolutely all of their qualia. Therefore, we should expect to have some qualia that we cannot introspect.

If all of our qualia were introspectible, then introspection would be an exceptional faculty. It may well be so. We do not, however, presently have any good reason to think that it should be. Until we find some particular reason to think that introspection should be remarkably comprehensive, the preceding reasoning suggests that it probably has limits. Consequently, we have reason to believe that at least some of our qualia are not introspectible.

3.2 The Argument from Varieties

The exceptionality argument centered on the likely limitations of introspection. The second argument that I will present focuses on the variety of kinds of hidden qualia that we might possibly have. Since it is independently plausible that we have many different varieties of qualia, it is *more than* merely plausible that we have at least one variety of hidden qualia. I have in mind a probabilistic argument: if we judge that there is a not too small probability that we have hidden qualia of each kind, then (given their independence) we are committed to thinking that it is fairly probable that we have at least one kind of hidden qualia. I will briefly describe five kinds of hidden qualia and present a few considerations for thinking that we might have hidden qualia of each kind. In my discussion of each, I will suggest that there are some *prima facie* considerations for thinking that we might have qualia of that kind, and good reasons to think that any such qualia would be hidden.

Remote Qualia

The human brain is roughly organized by cognitive sophistication: more sophisticated functions are handled by the cerebral cortex, and more basic functions are handled by the midbrain and hindbrain. While most philosophers and cognitive scientists associate our introspectible qualia with activities in the cortex, other qualia might also be produced in those regions of the midbrain and hindbrain that operate largely independently of the cortex.

Consider the cerebellum. The cerebellum is a large structure in the hindbrain that plays a substantial role in directing bodily action. The human cerebellum takes up approximately 10% of the volume of the brain and contains about half its total number of neurons. Absent a full understanding of exactly what neurons do to produce qualia, the simple fact that the cerebellum includes a large number of highly organized neurons that operate with the same basic mechanisms as neurons in the cerebral cortex gives us a weak reason to think that it too might produce qualia.

The case for cerebellar qualia is strengthened by the fact that one of the major functions played of the cerebellum is similar to one of the major functions of the cortex. The cerebellum helps to organize the movements of the muscles in the body. In this capacity, it receives and processes sensory information about the location of limbs, the tension in individual muscles, and the body's current state of balance and in turn, it contributes to the guidance of our actions.

Just like the cortex, the cerebellum takes in sensory information, processes it, and uses it to make decisions. It may not be as sophisticated as the human cortex, but in terms of complexity it is at least on par with the brains of simple vertebrates. For this reason, I think that the case for cerebellar qualia is about as strong as the case for qualia in the brains of relatively simple vertebrates. If we think that it is at least somewhat plausible that most vertebrates experience qualia because of the limited physical or functional similarities between their brains and ours, then we should think that it is at least somewhat plausible that we also have cerebellar qualia.

Conscious control or active monitoring over many regions of the brain outside of the cerebral cortex would serve little purpose, and as a result there are far fewer neural connections between the cognitive centers of the cortex and these other regions. Since introspection probably heavily involves the cognitive centers of the cortex, this means that even if these regions did contribute to our qualia, we should not expect to know about it. Given the absence of interaction between these regions and the cognitive centers of the cortex, any qualia produced remotely would be hidden.

Duplicate Qualia

Consciousness in human beings is typically attributed to processes in the cortex. The cerebral cortex consists in two separate hemispheres that are connected to each other primarily through a bridge of neurons called the ‘corpus callosum’. The hemispheres overlap significantly in the cognitive functions that they perform, and they are able to operate independently from each other (as we can see by the strikingly limited effects of severing the corpus callosum). The functional overlap of the hemispheres extends to sensory processing: each hemisphere contains its own visual, auditory, and somatosensory regions. Not only do the two hemispheres each process the same sorts of sensory information, but much of the same information is represented in the corresponding regions of the two cerebral hemispheres.⁸

It is extremely plausible that severing the corpus callosum results in the two hemispheres producing their own distinct qualia (Puccetti, 1981). If each hemisphere is capable of producing its own qualia when the corpus callosum is severed, it is reasonable to think that it also produces its own qualia even in an intact brain.

Since the qualia that we have are surely influenced by what information is represented in the sensory regions of our brain, and since much of the same information is represented in both

⁸Although some forms of sensory information received on one side of the body are first directed to the contralateral hemisphere, each hemisphere also receives information from the ipsilateral side of the body across the corpus callosum.

hemispheres of an intact brain, there is a moderately strong case that some of our qualia come in duplicate pairs. We cannot introspect the existence of distinct pairs of nearly identical qualia, so it is consequently fairly plausible that we have some qualia that are unintrospectible but that are otherwise much like qualia that we can introspect.

Insofar as the processes that are responsible for producing qualia are duplicated in the two hemispheres of our cortex, it is also plausible that we experience duplicate qualia. It would not provide any clear advantage to us to know about the duplicity of our qualia, so it makes sense that we should not be aware of some of our duplicate qualia.

Qualia Static

Radios reproduce sounds that are converted and transmitted through the air in the form of radio waves. Radios also produce some unintended sounds in the course of normal operation, as natural signals carried through the air interfere with a radio's intended purpose and cause it to produce a crackle.

It is natural to think that qualia are an end product of certain neural processes in something like the way that sounds are the end product of the operation of a radio's component parts. This raises the question: if sensory processes produce qualia similarly to the way that radios produce sounds as part of their normal operation, is there anything akin to radio static for qualia? *Qualia static* are qualia that are produced by disturbances in the neural processes that produce normal introspectible qualia.

It is well known that neural processes tend to be rather noisy. Neurons are prone to arbitrary firing even in the absence of external stimuli. The fact that the neurons associated with the processes that produce particular qualia are prone to arbitrary firing makes it plausible that we experience some kind of qualia static.

We may be aware of occasional disturbances produced by arbitrary neural activity, but we are not aware of pervasive qualia static. This, however, provides little evidence against pervasive qualia static, as there would be little value to knowing about them. A well-designed introspective faculty would tune out static to avoid cluttering our cognitive centers with useless information. Given what we know about the way the brain works and how it produces qualia, it is neither a priori implausible that we should have qualia static nor that we should lack introspective access to it.

Sleep Qualia

We commonly assume that sleep consists of, as Edgar Allen Poe put it, “little slices of death” that are devoid of any kind of conscious experience. While we have intermittent qualia in the form of dreams, we don’t dream continuously. If dream qualia are the only qualia we have while asleep, the common assumption is right.

Our brains do not shut down during sleep, nor do they behave exactly as they do while awake. There are widespread neurological changes in the kinds of activity of neurons throughout the cortex, but our neurons remain active throughout sleep. Since the neurons that produce qualia while the brain is awake continue to be active during sleep, they might continue to produce some kinds of qualia or other (perhaps even qualia static).

We don’t typically remember any qualia from the deepest stages of sleep. Whatever we do experience during these stages must not make an impression on our memory. However, this provides little evidence that we in fact experience nothing. Sleep is known to have a significant effect on memory, and even dreams are often quickly forgotten.

If we do have some kinds of qualia during the deepest stages of sleep, it is not implausible that we should be unable to introspect them. The deficits of memory involved in sleep provide us with reason to think that some higher-level faculties are taken offline. It would not be surprising to find out that the same was true for the mechanisms involved in introspection as was true for the mechanisms involved in memory. If sleep impedes our ability to introspect, then we might have hidden qualia throughout its deepest stages.

Constant Qualia

On a standard view of the mind, our experiences are fleeting: one set of qualia is continually being replaced by another and no particular quale lasts long outside of very specific contexts. On the other hand, there might be some qualia that are with us perpetually. Such qualia might arise from the basic operations of the brain or as a byproduct of the functioning of our sensory systems.

Chalmers (Chalmers, 1997, 10) speaks of one possible variety of such qualia when he describes a qualitative *sense of self* as a “background hum” that is fundamental to consciousness. If Chalmers is right about the existence of this sort of constant quale, then it is natural to wonder whether there might be others as well.

A priori, we have no reason to rule out the possibility that we should have some kinds of constant qualia beyond Chalmers’ sense of self. Considerations of parsimony might count against

the postulation of additional constant qualia, but such considerations are typically fairly weak. If we have good reasons to rule out constant qualia, they must come from our experiences.

We cannot introspect any qualia that remain with us continuously (other than Chalmers' sense of self). This is consistent with the existence of constant qualia. The capacity to introspect constant qualia would not be especially useful. Being more or less unchanging, such qualia could never provide us with novel reasons to act in any particular ways. Consequently, the fact that we cannot introspect constant qualia gives us little reason to reject their existence.

The five possible kinds of hidden qualia that I have surveyed are not exhaustive. The human brain is extremely complicated. It contains a large amount of neurological variety and carries out a huge number of different functions. There are surely many other epistemically possible sources of hidden qualia. Given that we can find a number of different varieties of hidden qualia that are independently plausible, we can conclude that hidden qualia are fairly probable.

4 Justification and the Absence of Qualia

I hold that in light of the exceptionality argument and the argument from varieties, it is unreasonable to believe that we lack hidden qualia. In this final section, I will explore whether we might borrow a strategy from the literature on epiphenomenal qualia and treat our acquaintance with our qualia as justification enough for believing in the non-existence of hidden qualia *on the assumption that we do lack them*. Such a result would be of interest even though its significance is somewhat limited. It would provide those gripped by the possibility of hidden qualia with no reason to change their minds. However, if it could be established, then hidden qualia's doubters could at least feel comfortable in their disbelief.

The epistemology of hidden qualia shares much with the epistemology of epiphenomenal qualia. Epiphenomenal qualia are qualia that have no causal influence over the physical states of our brains. Many have found epiphenomenalism to be deeply attractive because it allows us to reconcile anti-physicalist intuitions with belief in the causal closure of the physical world. However, it is hard to see how we could ever know about epiphenomenal qualia. Assuming that our beliefs are determined by our brain states, our belief in qualia can't be *sensitive* to the existence of epiphenomenal qualia – given how we're physically constituted, our beliefs about our qualia can't

depend on our actually experiencing any epiphenomenal qualia. Hidden qualia are epistemically like epiphenomenal qualia insofar as neither have any direct causal impact on our beliefs about our qualia.

Opponents of epiphenomenalism (such as Moore 2012) have suggested that the lack of causal influence should force epiphenomenalists into radical skepticism about their own qualia. Like epiphenomenal qualia, hidden qualia do not influence our higher-order beliefs. While our armchair epistemic position with respect to hidden qualia is comparable to our position with respect to epiphenomenal qualia, the challenge that they pose is different. We tend to assume that while we have introspectible qualia, we lack hidden qualia. The epistemic challenge for the epiphenomenalist is to explain how it is that we can be justified in believing in the *existence* of epiphenomenal qualia. In contrast, the epistemic challenge arising from hidden qualia is how we can be justified in believing that they *don't exist*.

The qualia we have constitute what it is like to be us. Our qualia are our access to the world – they are that part of reality that we are most intimately familiar with. David Chalmers' initial response to the epistemic challenge for the epiphenomenalist was to suggest that the intimate relation we have with our qualia precludes the need for any kind of causal influence over our belief states. The virtue of intimacy outweighs the vice of insensitivity.

Chalmers explains:

Conscious experience lies at the center of our epistemic universe... it is *having* the experiences that justifies the beliefs. For example, the very fact that I am having a red experience now provides justification for my belief that I am having a red experience. (Chalmers, 1997, 196).

I will use the term 'acquaintance' for the intimate relation we have with our qualia by virtue of the fact that they are phenomenal characters. Chalmers did not specify the precise limits of his proposal, but it might be read to say that *any* correct beliefs about our qualia are justified by the fact that we are acquainted with them. Call this the 'liberal intimacy hypothesis'.

What it is like to be us is determined in part by what qualia we lack. If we lack hidden qualia, then the liberal intimacy hypothesis implies that we are justified in disbelieving in them. By the same token, if we have hidden qualia, then we are justified in believing in them. Suppose that you do in fact experience sixteen separate kinds of constant qualia and that you come to believe this

without introspecting them – say, because you like the number sixteen – then the liberal intimacy hypothesis would suggest that your belief would be justified.

Ernest Sosa (2003) presented a persuasive argument against this idea. According to Sosa, the liberal intimacy hypothesis predicts that we will have justification in cases where we intuitively don't. This is a problem best seen through examples. Take a few seconds to examine the matrix of gray dots.

There is no dot of the shade ■ in the figure. If the liberal intimacy hypothesis were right, then we would be *prima facie* justified by virtue of our acquaintance with our qualia in believing that such a dot is not in the matrix. This is implausible. To be justified in believing that there is no such dot, we must take the time to carefully survey the figure. We are not justified in believing *just anything* that happens to be true about our qualia. Something beyond intimacy is necessary.

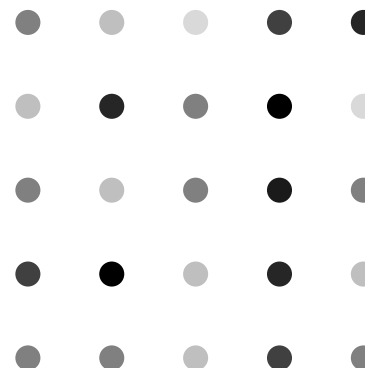


Figure 1: dots

If something more than merely having experience is necessary for believing that one has an experience of that sort, then an absence of hidden qualia need not by itself justify disbelief in hidden qualia. In the wake of Sosa's challenge, a variety of proposals have been suggested for what this something more might be. Sosa proposed that we are justified in believing the products of a reliable introspective mechanism. Richard Feldman (2004) proposed that the justification of a belief that we are having a particular sort of experience requires not only that we have a given experience, but that we attend to its relevant features. Richard Fumerton (2005) suggested that we must not simply have a quale, but we must have some special epistemic connection with it (which Fumerton calls 'direct acquaintance') that we don't have with all of our qualia. Ted Poston (2007) drew upon the details of Fumerton (1995) to suggest that justification requires being directly acquainted with the correspondence between a quale and our belief about it.

None of these alternatives will vindicate a belief in the absence of hidden qualia. Sosa's reliabilist justification has been aptly criticized by Feldman (2004) and Poston (2007) on the grounds that accidental reliability doesn't provide justification. Sosa's view predicts that a person in a world of

hens who all have forty-eight speckles might be justified in believing that a particular hen-experience involved forty-eight-speckles, even if the judgment were not responsive to the cardinality of the hen's perceived speckles. The problem seems to be, in part, that the reliability of the belief is accidental. It is similarly dubious that we are justified in disbelieving in hidden qualia simply because our disbelief is accidentally reliable. If it is a matter of sheer evolutionary luck that we correctly and reliably disbelieve in hidden qualia, we aren't justified.

Each of the other three accounts suggests that justification requires some sort of intellectual focus on the relevant aspects of one's experience. Feldman took this to be attention, and Fumerton and Sosa to be direct acquaintance. Though direct acquaintance is a bit mysterious, it is plausibly very closely related to attention; we are not directly acquainted with qualia (or any correspondence between them and our beliefs) to which we cannot attend. Furthermore, it is plausible that we are unable to attend to the absence of hidden qualia, and so we are not justified in believing that they don't exist.

Cognitive theories of attention suggest that attention involves shifting our neural resources so that some cognitive processes have better access to important cognitive channels. By gaining access to those channels, they can exert a greater influence over our higher-order intellectual faculties. They can come to play a greater role in our deliberations or enter into our memories. We attend to visual experiences by strengthening the effect that visual processing regions have on our higher-order faculties and we attend to what we hear by strengthening the effect that the auditory processing regions have.

If this is correct, then it seems likely that it will be impossible for us to focus on any qualia *or their absence* that we are not able to access introspectively. For instance, supposing that there really are no pathways between the cerebellum and belief-generating parts of the cortex, then the normal mechanisms for shifting neural resources may do nothing to grant us better intellectual access to what goes on in the cerebellum. It is plausible that we cannot focus on the activities of the cerebellum in order to attend to, or become directly acquainted with, the absence of cerebellar qualia, and also that we can't attend to or be acquainted with duplicate qualia, qualia static, or constant qualia.

Chalmers (2003) has expanded and refined his original proposal. The new version traces our justification for our beliefs about qualia to a special class of fundamental beliefs – *direct phenomenal*

beliefs – which predicate “the concept of the very experience responsible for constituting its content.”

(236) The idea is that certain kinds of concepts refer by way of those concepts’ constituents. A direct phenomenal concept includes the very experience of redness to which it refers. Direct phenomenal beliefs apply a direct phenomenal concept to the very instance of the property that the concept includes.

This proposal neatly dodges any worries about insensitivity because it provides a way for epiphenomenal qualia to influence beliefs (by supplying part of their contents) without influencing physical states. It also gives attention a role to play, because justification requires forming beliefs that include the very qualia that we’re experiencing. Interestingly, such beliefs are incorrigible because the tokened quale determines the concept’s reference in a way that precludes a direct phenomenal belief from ever being wrong. Nevertheless, Chalmers suggests that such beliefs are still justified by their intimacy, rather than their incorrigibility.⁹ Chalmers’ new proposal gives acquaintance a more moderate role: we are not justified in believing anything about what we are acquainted with, we are specially justified in direct phenomenal beliefs by the form of those beliefs together with the facts about acquaintance.

Chalmers’ proposal says that our knowledge of our qualia depends for its justification on inferences from direct phenomenal beliefs. Direct phenomenal beliefs can give us no grounds for belief in the non-existence of hidden qualia. We can’t form direct phenomenal beliefs about the non-existence of qualia. Nor can we infer the non-existence of hidden qualia from any positive beliefs about the existence of qualia. Chalmers doesn’t directly address this issue, but his remarks elsewhere suggest that he might accept this conclusion.¹⁰ It is a surprising result because it suggests that our beliefs about the kinds of qualia that we lack are justified in a fundamentally different way, if they are justified at all, from our beliefs about what qualia we have.

No matter what their merits for justifying belief in epiphenomenal qualia, neither Chalmers’ original proposal nor its later refinement provides us with a solid justification for belief that we lack hidden qualia, even if we don’t have any.

⁹This is because we cannot know whether a given direct phenomenal belief is a direct phenomenal belief. Mathematical beliefs are often incorrigible, but that doesn’t justify them all by itself.

¹⁰Here is Chalmers on the incorrigibility of disbelief: “Incorrigibility theses are also sometimes articulated in a negative form, requiring that a subject cannot be mistaken in their belief that they are *not* having a given sort of experience. No direct phenomenal belief is a negative phenomenal belief, so the current framework does not support this thesis, and I think the thesis is false in general.” (2003, 245)

5 Conclusion

One intuitive view of the mind holds that qualia are real, that they are totally distinct from our beliefs about our experiences, and that they can't evade our introspective faculties. I hope to have shown that there is a tension between these ideas. It is reasonable to reject one to save the others. If we accept the idea that we have qualia that are distinct from our beliefs about our qualia, we should be willing to accept that hidden qualia are not only possible, but actually somewhat likely.

Though some may recoil at this proposal, I don't think they should. While hidden qualia are deeply counter-intuitive, there is nothing incoherent about them. The world is a surprising place, and our intuitive views about it have turned out to be wrong in a number of unforeseeable ways. We have already discovered much about our minds that we did not expect to. The possibility that we are radically mistaken about the extent of our current conscious qualia should not be dismissed too quickly.

References

- Ned Block. On a Confusion About a Function of Consciousness. *Behavioral and Brain Sciences*, 18(2):227–247, 1995.
- Ned Block. Consciousness, Accessibility, and the Mesh Between Psychology and Neuroscience. *Behavioral and Brain Sciences*, 30(5):481–499, 2007.
- Ned Block. Consciousness and Cognitive Access. *Proceedings of the Aristotelian Society*, 108(3): 289–317, 2008.
- Alex Byrne. Some Like It HOT: Consciousness and Higher-Order Thoughts. *Philosophical Studies*, 2(2):103–29, 1997.
- Peter Carruthers. Brute Experience. *Journal of Philosophy*, 86:258–269, 1989.
- Peter Carruthers. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge University Press, 2003.
- David Chalmers. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, 1997.

- David Chalmers. The Content and Epistemology of Phenomenal Belief. In Q. Smith and A. Jokic, editors, *Consciousness: New Philosophical Perspectives*. Oxford University Press, 2003.
- Daniel Dennett. Quining Qualia. In A. Marcel and E. Bisiach, editors, *Consciousness in Modern Science*. Oxford University Press, 1988.
- Richard Feldman. The Justification of Introspective Belief. In Earl Conee and Richard Feldman, editors, *Evidentialism: Essays in Epistemology*. Oxford University Press, 2004.
- Richard Fumerton. *Metaepistemology and Skepticism*. Rowman & Littlefield Publishers, 1995.
- Richard Fumerton. Speckled Hens and Objects of Acquaintance. *Philosophical Perspectives*, 19(1): 121–138, 2005.
- Michael Graziano. *Consciousness and the Social Brain*. Oxford University Press, 2013.
- Christopher Hill and Brian McLaughlin. There are Fewer Things in Reality Than are Dreamt of in Chalmers’s Philosophy. *Philosophy and Phenomenological Research*, 59(2):445–454, 1999.
- Dwayne Moore. On Robinson’s Response to the Self-Stultifying Objection. *Review of Philosophy and Psychology*, 3(4):627–641, 2012.
- David Papineau. *Thinking About Consciousness*. Oxford University Press, 2002.
- Ted Poston. Acquaintance and the Problem of the Speckled Hen. *Philosophical Studies*, 132(2): 331–346, 2007.
- Jesse Prinz. A Neurofunctional Theory of Visual Consciousness. *Consciousness and Cognition*, 9(2):243–259, 2000.
- Roland Puccetti. The Case for Mental Duality: Evidence from Split-Brain Data and Other Considerations. *Behavioral and Brain Sciences*, 4:93–99, 1981.
- David Rosenthal. *A Theory of Consciousness*. Zentrum für Interdisziplinäre Forschung, 1990.
- Eric Schwitzgebel. Do You Have Constant Tactile Experience of Your Feet in Your Shoes?: Or Is Experience Limited to What’s in Attention? *Journal of Consciousness Studies*, 14(3):5–35, 2007.
- Eric Schwitzgebel. *Perplexities of Consciousness*. MIT Press, 2011.

Ernest Sosa. Privileged Access. In Quentin Smith and Aleksandar Jokić, editors, *Consciousness: New Philosophical Perspectives*. Cambridge University Press, 2003.

Michael Tye. *Ten Problems of Consciousness: A Representational Theory of The Phenomenal Mind*. MIT Press, 1995.

Jessica Wilson. What Is Hume's Dictum, and Why Believe It? *Philosophy and Phenomenological Research*, 80(3):595–637, 2010.