

Explaining Qualia

Kiefer Owens
University of Illinois at Urbana-Champaign

ABSTRACT

In this paper, I consider what the data of qualia imply about which theory of consciousness is true. I argue that no physicalist alternative accounts both for their ineffable, experiential nature and their causal efficacy. Substance dualism, on the other hand, does account for both of these. Qualia thus constitute evidence in favor of dualism over and against its physicalist competitors, and this shows that it is a live option in philosophy of mind.

Key words: qualia, consciousness, physicalism, reductive physicalism, epiphenomenalism, substance dualism, functionalism, topic-neutral, ineffable, causal efficacy

INTRODUCTION

One of the goals of theories about consciousness is to explain qualia, or the sensory qualities of some mental phenomena, such as the hurtiness of pain or the appearance of green. In other words, the primary views in philosophy of mind are distinguished in part by how they fit qualia into their theories of consciousness. How successful a theory is in explaining these data of qualia is often seen as a major factor in determining its plausibility. I am assuming here that our basic experience of qualia constitute data to be explained by our theories of consciousness; that is, people have a natural, intuitive grasp of qualia (even if they do not express this in a philosophical way), and these must be made sense of by our theories in a way that does justice to our intuitions. This line of thinking has been challenged by some in experimental philosophy, where it is argued that intuitions about qualia are not so basic and uniform. These considerations go beyond the scope of this paper and will not be represented here. In this paper, I will explain and assess the major theories in philosophy of mind regarding qualia—substance dualism, reductive physicalism, and epiphenomenalism. I argue that only substance dualism explains

or incorporates both the ineffable, experiential nature of qualia (i.e., their raw feel) and their causal efficacy.

My motivations for writing this paper rest in the belief that substance dualism still has something to offer the philosophy of mind, and therefore should not be dismissed out of hand as a theory of the past. As a naturalized philosophy of mind develops, we continue to run into cases where our interpretation of what science says about the world conflicts with our ordinary or “common sense” conception of the world. In many cases this is a good thing; science is often in a better position to know about some things than common sense. But I wonder where the line is with this project. That is, how steadfast should we be with our desire to construct naturalist or physicalist pictures of the world if they continue to deny the basic intuitions we share? Rather, if we have alternative theories that preserve these intuitions and do not themselves suffer from any insuperable difficulties, then is it reasonable to give these theories consideration? I maintain that it is, and intend in this paper to show that the debate about qualia is one area where an alternative theory like dualism should be reconsidered. Later in the paper, I also show that the difficulties about dualism are not unresolvable. Thus, since it preserves our intuitions about qualia, it should again be considered a live option in philosophy of mind. Before we see this, however, we must be clear on what each theory says about qualia.

THEORIES OF QUALIA

According to René Descartes, qualia demonstrate that the mind is distinct from the body. He argues as follows. I am distinctly aware of myself as a thinking thing, which means in part that I am a thing that has sensory perceptions or qualia (e.g., I feel pain). But I am not distinctly aware of myself as having a body; I can conceive of myself existing and having qualia without a body. Hence, it is possible for me to exist without my body, from which it follows that I am an immaterial substance distinct from my body. Nevertheless, it is clear that I and my body interact with each other (Descartes 1991, 23–9). On Descartes’ substance dualism, then, qualia are states of the immaterial soul that stand in causal relations with the body.¹

However, as Jaegwon Kim points out, dualism faces the problem of explaining how immaterial minds could causally influence the physical world, since physical

1. In a more technical sense, it is the immaterial soul itself that has causal relations with the body, and the qualia are properties or features of the soul. However, qualia would still play a part in the overall causal connection between the soul and the body (e.g., a feeling of pain is part of the reason why the soul causes the body to wince).

effects plausibly require physical causes (Kim 2006, 42).² In fact, Kim argues, the problem of mental causation shows that qualia can only be explained within reductive physicalist or epiphenomenalist frameworks (Kim 2006, 194–200). For, given physicalism, we accept that every event with a cause has a physical cause (i.e., physical causal closure), and that a person's mental states are fixed by their physical states (i.e., mind-body supervenience). But then the only way to say that qualia have physical effects is to say that qualia reduce to brain states (reductive physicalism), since physical effects already have physical causes, and we do not want to admit overdetermination.³ In other words, given that every brain state has a sufficient physical cause, there is no room left for the causal efficacy of mental states *unless* they are just the same as brain states.⁴ On the other hand, if we are not concerned with preserving mental causation, we may say that qualia are causally inert immaterial properties that supervene on physical properties (epiphenomenalism). Kim himself finds the existence of mental causation to be non-negotiable and therefore takes the reductive physicalist route, though he acknowledges that both views have their own problems (Kim 2006, 181).

The reductive physicalist option may be represented by J.J.C. Smart. His primary reason for believing that consciousness is physical is that everything else about the world is explicable by physical science, and it is absurd to think that consciousness is the one exception (Smart 1991, 169–70). What then does he do with qualia? These, Smart argues, are not sets of the intrinsic properties of consciousness but rather vague apprehensions of what are in fact brain states (Smart 1991, 172–3). Thus, to say, "I have a green sensation," is merely to say something like: "I am in the kind of state I would be in if I were seeing something green." This explains both how qualia can be identical with brain states and how they may retain their elusive nature. Reports of qualia are therefore vague, topic-neutral descriptions of the brain states to which they reduce.

The epiphenomenalist option may be represented by Frank Jackson. He has three reasons supporting the non-physicality of qualia (Jackson 1982, 128–32). First, we can have complete physical knowledge about certain experiences and

2. This problem was initially raised by Princess Elisabeth of Bohemia in response to Descartes, and is seen by many as the fatal flaw of substance dualism.

3. Overdetermination occurs when there is a sufficient cause for some effect that already has a sufficient cause at the same time. As Kim points out, these cases are rare and should not be invoked lightly.

4. The same argument can also be used to show that qualia do not cause other *mental* states either.

compos mentis

still lack knowledge of those experiences: that is, knowledge of the qualia. Second, it is possible for there to be organisms exactly like us in every physical respect, but different from us in that they lack a conscious mental life. Third, following Nagel, physical information alone cannot tell us *what it is like* to have an unfamiliar perspective on the world (e.g., what it is like to be a bat). All of these considerations aim to show that all the physical information about us leaves out some information: namely, the non-physical qualia. Nevertheless, Jackson holds, there is no good reason to think that qualia have any influence on the physical world, which leads to his epiphenomenalist position (Jackson 1982, 133–5). He admits that it remains mysterious why qualia should exist and be caused by brain states, but says that we should not expect to understand this due to our natural epistemic limitations.

ASSESSMENT: ALTERNATIVES TO DUALISM

Kim's exclusion argument. It seems to me that Kim is right in arguing that reductive physicalism and epiphenomenalism are the only options for explaining qualia once dualism is denied. If dualism is false, then—granted we are not idealists either—we must hold that consciousness is in some sense grounded in a physical substance like the brain. But if that is the case, then qualia will either be identical with brain states, or be properties or features realized by brain states. Moreover, given this kind of physicalism, we are apt to accept Kim's premises of physical causal closure and mind-body supervenience. But if that is the case, then qualia can only be causally efficacious if they are brain states—otherwise they are excluded. So since they either are or are not causally efficacious, either reductionism or epiphenomenalism is true.

Some may worry that this unjustifiably leaves out functionalism, which is prominent view in philosophy of mind. On this view, mental states (and hence qualia) are functional events of an organism; that is, an organism's consciousness is simply an organization of its functions (Putnam 1991, 199–200).⁵ Functionalism thus portrays mental states in a topic-neutral way, saying nothing about their intrinsic nature. But this is precisely why functionalism is left out: it may tell us something informative about qualia, but it does not tell us anything about what qualia *are in themselves*, and therefore has no explanation of them. Notably, functionalism does not even tell us if qualia are physical or non-physical, since it is consistent

5. For example, pain would be a state that is caused by damage to the body and typically causes avoidance behavior, wincing, etc.

with both dualist and monist pictures of consciousness (i.e., an immaterial soul may be part of an organism's functional organization). In fact, functionalism's lack of an explanation of qualia constitutes one of the biggest objections to the view, for many feel that it is possible for there to be organisms that have the same functional organization as us, but that have absent or inverted qualia (Kim 2006, 162–3). All this goes to show that functionalism is not a live option for explaining qualia. Kim's argument thus succeeds in showing reductive physicalism and epiphenomenalism to be the alternatives to dualism in explaining qualia.

Smart's reductive physicalism. How does Smart's reductionism fare? Recall that his primary reason given for endorsing this view is that allowing non-physical consciousness would threaten the unity and explanatory power of science, and Smart simply cannot accept this. Nevertheless, he himself admits that this is "largely a confession of faith," not amounting to much of an argument (Smart 1991, 170). I think he is right here: it would be question-begging to assume that science must explain everything, and then conclude that consciousness is therefore physical, since this conclusion is *equivalent* to saying that consciousness is explicable by science. In other words, the conclusion that science explains everything should come *after* we determine whether reductionism is true, not before it. So there is not much justification to be gained from this end. Smart's view does have something going for it in that it allows for qualia to be causally efficacious, as Kim's argument showed.

But reductionism runs into an insuperable difficulty: by reducing qualia to brain states, the ineffable, experiential nature of qualia is removed. The argument for this is simple: any two identical things must have all the same properties, but qualia clearly have properties that brain states do not and vice versa. Consider the appearance of green. If this quale was a brain state, then it would make sense to ascribe to it a certain weight, location in space (say, closer to the left ear than the right), and capacity of being seen by a neurophysiologist (Moreland and Craig 2003, 233). These are all properties of brain states, but it seems bizarre to ascribe them to qualia; hence, qualia are not brain states. Jackson's arguments demonstrate the same point. Consider just one: the case of the neurophysiologist Mary (Jackson 1982, 130). Mary specializes in the neurophysiology of vision. She possesses all the scientific knowledge there is to know about what goes on in our brain when we experience color, despite working in a black and white room all her life and only possessing a black and white television. Nevertheless, when she leaves the room and experiences green for the first time, we still want to say that

Mary gains knowledge about the world. Thus, qualia are not included in all the information about brain states; they have different properties and are therefore not identical. These arguments serve to show that qualia cannot be reduced to brain states without sacrificing the very thing that makes them qualia: their ineffable, experiential nature.

Smart does attempt to preserve the ineffable nature of qualia with his topic-neutral approach (i.e., qualia are not things or properties in themselves but vague apprehensions of brain states). However, this attempt fails for the simple reason that, contrary to Smart, people do not think of qualia in this way. When people say "I have a green sensation," they are not vaguely referring to some *other* event that often happens, but rather to their *present* experience of having an appearance of green. Reductive physicalism thus fails to adequately explain qualia because it removes their ineffable, experiential nature.

Jackson's epiphenomenalism. But if reductionism fails, what about epiphenomenalism? By denying the physicality of qualia, Jackson does seem to preserve their experiential nature. However, as we saw earlier, this view does not allow for their causal efficacy, a tenet that most of us take to be non-negotiable. As Jerry Fodor explains:

[I]f it isn't literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching, and my believing is causally responsible for my saying..., if none of that is literally true, then practically everything I believe about anything is false and it's the end of the world. (cited in Kim 2006, 181)

This is a perhaps overstated way to make the point that mental causation is a basic intuition that cannot reasonably be denied. But then views like epiphenomenalism must be deemed false.

Jackson responds to this point by raising the possibility that our intuition is mistaken. He notes:

No matter how often *B* follows *A*, and no matter how initially obvious the causality of the connection seems, the hypothesis that *A* causes *B* can be overturned by an over-arching theory which shows the two as distinct effects of a common underlying causal process. (Jackson 1982, 133)

This may be true, but is hardly an adequate response as it unjustifiably shifts the burden of proof. If mental causation seems obvious to us (as Jackson seems to concede), then we are perfectly within our rational rights to believe in it in the absence of some overriding defeater. However, the possibility that mental causation does not happen after all no more defeats our belief in it than the possibility of us being in the matrix defeats our belief in the external world. Jackson is right that he has an over-arching theory that would expose the illusion of mental causation, but we have strong reason to disbelieve that theory in light of our basic intuition that mental causation does in fact occur. So denying the causal efficacy of qualia constitutes a failure on the part of epiphenomenalism to adequately explain a necessary feature of qualia, and we may therefore conclude that neither physicalist alternative is a plausible account of qualia.

RECONSIDERING DUALISM

If neither physicalist account can explain qualia, what about substance dualism? This view, like epiphenomenalism, posits qualia as immaterial features, which preserves their ineffable, experiential nature. Moreover, it also holds that qualia are causally efficacious, doing justice to that intuition as well.

We recall, however, that the primary problem with dualism was showing just how an immaterial entity like the soul and a physical entity like the body could interact. This problem may take different forms. It may take the form of a question (e.g., "How could immaterial qualia affect a physical body?") that demands answering before dualism should be believed. This is the version that Elizabeth seemed to espouse in her letter to Descartes (Kim 2006, 41-2). However, as Descartes responded, we have a basic apprehension of the interaction between our minds and bodies, such that we know *that* it occurs even without having an explanation of *how* it occurs (Descartes 1991, 33). That is, the objection assumes that we must know *how* a causal interaction occurs in order to know *that* it occurs, but this is unjustified (Moreland and Craig 2003, 243).⁶ Moreover, as J. P. Moreland and William Lane Craig explain, such a "how" question assumes that some intermediate mechanism be given by which the soul influences the body (Moreland and Craig 2003, 243-4). But the interaction between the soul and the body may well be immediate, having no such mechanism. Therefore, this version of the interaction

6. As these authors argue, if this assumption were true, then we could not know that causal forces such as gravity exist since we do not have an account of how gravitational interaction occurs.

problem does not constitute a difficulty for dualism.

A more sophisticated form that this objection takes is based off the notion of physical causal closure and runs as follows (Kim 2006, 42; Robinson 2012, sec. 3.1; Collins 2008, sec. I and II). A basic principle in modern physics is the conservation of energy and momentum, which states that, in closed or causally isolated physical systems, the total amount of energy and momentum must remain constant. However, dualist interactionism would violate this law, since the mind must produce new energy and momentum in the brain to create a new neural event. Thus, such interaction cannot take place. Two points may be made by means of response. First, this argument begs the question, since applying the conservation law to the body just is to deny that dualism is true. That is, the dualist denies that the body is a closed system, so the conservation law becomes irrelevant in this case and therefore is not violated (Robinson 2012, sec. 3.1). Second, as Robin Collins notes, the conservation principle is not a universal principle in physics:

The [energy conservation] objection against interactionistic dualism fails when one considers the fact that energy conservation is not a universally applicable principle in physics [e.g., general relativity] and that quantum mechanics sets a precedent for interaction (or at least law-like correlation) without any sort of energy-momentum exchange, or even any intermediate carrier. (Collins 2008, sec. V)

But if this principle does not apply in some areas of physics, why think that it must apply in the case of mind-body interaction? If it does not even apply to the universe as a whole, then it is not necessarily true that it applies to the brain, either. Both versions of the interaction problem therefore fail.⁷

So given the options available to us for explaining qualia, substance dualism commends itself as the most plausible in that it explains both their ineffable,

7. Another objection to substance dualism is that it is incoherent: the very idea of an "immaterial substance" is contradictory since the idea of a "substance" is of a physical thing. But this objection plainly begs the question against dualism, which is built on arguments that show that there are immaterial substances. Dualists (and theists, for that matter) have the idea of a substance that is immaterial, and use the term "substance" merely to pick out a thing which has properties, whether it is physical or non-physical. Within a physicalist framework, it is natural to be perplexed at the idea of an immaterial substance, but the very question under consideration is whether a physicalist framework is the right framework. So this objection is a non-starter.

experiential nature and their causal efficacy, where the physicalist alternatives fail in one of these respects.⁸

CONCLUSION

In this paper, we have examined different positions in philosophy of mind when it comes to explaining qualia, or the raw feels of experience. Substance dualism portrays them as features of the immaterial soul that causally interacts with the body. Reductive physicalism identifies them as nothing more than vague apprehensions of brain states. Epiphenomenalism conceives of them as non-physical properties that supervene on brain states but do not cause anything. Kim's explanatory exclusion argument was given to restrict the physicalist options to just reductionism and epiphenomenalism once dualism is denied. However, we saw that neither option adequately explains qualia: reductionism fails to fit in to its account their ineffable, experiential nature, whereas epiphenomenalism denies their causal efficacy; both moves are unacceptable. Dualism, on the other hand, explains both of these features of qualia without having any insuperable difficulties. It therefore commends itself as the most plausible theory.

The conclusions of this paper constitute an argument for substance dualism. If qualia cannot be incorporated into a physicalist picture of consciousness, then they count as evidence in favor of non-physicalism (i.e., dualism). Of course, other non-physicalist accounts like idealism were not considered here; it is up to the reader to examine how substance dualism fares against these accounts. Moreover, other considerations besides those about qualia may factor into one's overall assessment of these positions, both physicalist and non-physicalist. Nevertheless, such an argument does show that qualia constitute a defeater for physicalist theories of consciousness, which is not insignificant. Hence, the debate about qualia demonstrates the benefit of reconsidering substance dualism as a live option in contemporary philosophy of mind.⁹

8. A further question we might have concerns how it is the qualia function on substance dualism. Constructing a detailed substance dualist account of the experience of qualia, as well as of the human person as a whole, could either enhance or weaken the plausibility of the theory. I leave this project aside as an idea that deserves future consideration.

9. Thanks to Jonathan A. Waskan, some of the faculty of the philosophy department at the University of Michigan-Flint, and my undergraduate peers at the Michigan Undergraduate Philosophy Conference (held March 16, 2013) for helpful comments and suggestions on an earlier draft of this paper.

compos mentis

REFERENCES

Collins, Robin. 2008. "Modern Physics and the Energy Conservation Objection to Mind-Body Dualism" [condensed Version]. *The American Philosophical Quarterly* 45 1: 31–42.

Descartes, René. 1991. "René Descartes, Selections." In *The Nature of Mind*, edited by David M. Rosenthal, 21–37. New York and Oxford: Oxford UP.

Jackson, Frank. 1982. "Epiphenomenal Qualia." *The Philosophical Quarterly* 32.127: 127–36.

Kim, Jaegwon. 2006. *Philosophy of Mind* (2nd ed.). United States of America: Westview.

Moreland, J.P., and William Lane Craig. 2003. *Philosophical Foundations for a Christian Worldview*. Downers Grove, Illinois: IVP Academic.

Putnam, Hilary. 1991. "The Nature of Mental States." In *The Nature of Mind*, edited by David M. Rosenthal, 197–203. New York and Oxford: Oxford UP.

Robinson, Howard. 2012. "Dualism." In *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), edited by Edward N. Zalta.

Smart, J.J.C. 1991. "Sensations and Brain Processes." In *The Nature of Mind*, edited by David M. Rosenthal, 169–76. New York and Oxford: Oxford UP.