Qualia

Qualia are the subjective or qualitative properties of experiences. What it feels like, experientially, to see a red rose is different from what it feels like to see a yellow rose. Likewise for hearing a musical note played by a piano and hearing the same musical note played by a tuba. The qualia of these experiences are what give each of them its characteristic "feel" and also what distinguish them from one another. Qualia have traditionally been thought to be intrinsic qualities of experience that are directly available to introspection. However, some philosophers offer theories of qualia that deny one or both of those features.

The term "qualia" (singular: quale and pronounced "kwol-ay") was introduced into the philosophical literature in its contemporary sense in 1929 by C. I. Lewis in a discussion of sense-data theory. As Lewis used the term, qualia were properties of sense-data themselves. In contemporary usage, the term has been broadened to refer more generally to properties of experience. Paradigm examples of experiences with qualia are perceptual experiences (including nonveridical perceptual experiences like hallucinations) and bodily sensations (such as pain, hunger, and itching). Emotions (like anger, envy, or fear) and moods (like euphoria, ennui, or anxiety) are also usually taken to have qualitative aspects.

Qualia are often referred to as the phenomenal properties of experience, and experiences that have qualia are referred to as being phenomenally conscious. Phenomenal consciousness is often contrasted with intentionality (that is, the representational aspects of mental states). Some mental states—for example, perceptual experiences—clearly have both phenomenal and intentional aspects. My visual experience of a peach on the kitchen counter represents the peach and also has an experiential feel. Less clear is whether all phenomenal states also have intentional aspects and whether all intentional states also have phenomenal aspects. Is there really something that it is like to have the belief—even the occurrent belief—that there is a peach on the counter? What could be the representational content of the experience of an orgasm? Along these lines, the nature of the relationship between phenomenal consciousness and intentionality has recently generated considerable philosophical discussion. Some philosophers think that phenomenal consciousness reduces to intentional content, while others think that the reductive relationship goes in the other direction. Still other philosophers deny both claims.

From the standpoint of introspection, the existence of qualia seems indisputable. It has, however, proved remarkably difficult to accommodate qualia within a physicalist account of the mind. Many philosophers have argued that qualia cannot be identified with or reduced to anything physical, and that any attempted explanation of the world in solely physicalist terms would leave qualia out. Thus, over the last several decades, qualia have been the source of considerable controversy in philosophy of mind.

Table of Contents

1. The Hard Problem of Consciousness
2. Qualia and Functionalism
1. The Hard Problem of Consciousness

One of the most fundamental questions about the mind concerns its relationship to the body (and, more specifically, its relationship to the brain). This has become known as the mind-body problem. Although it dates back at least to Plato’s *Phaedo*, the problem was thrust into philosophical prominence by René Descartes. In taking up these issues in his *Meditations on First Philosophy*, Descartes argued for a dualist view according to which the mind and the body are fundamentally different kinds of things: While the body is a material thing existing in space, the mind is an immaterial thing, one that altogether lacks spatial extension. In contrast to dualists, the materialists claim that everything that exists must be made of matter. Historically, materialism was associated with Thomas Hobbes. Starting in the twentieth century, this position has become known as physicalism, the claim that everything that exists—all things and all properties of things—must fundamentally be physical. Most philosophers today endorse some form of physicalism.

For some aspects of consciousness, it is relatively straightforward to see how they can be accommodated within a physicalist picture. Consider, for example, our abilities to access, report on, and attend to our own mental states. It seems reasonable to assume that as neuroscience progresses and we learn more and more about the brain, we will be able to explain these abilities in terms of neural mechanisms. Aspects of consciousness that can be explained in this way constitute what David Chalmers has referred to as the easy problems of consciousness. The assertion that these problems are easy does not mean that they have already been solved or even that we are close to finding solutions. As Chalmers explicitly notes, we should think of “easy” as a relative term. In most cases, we are still nowhere near having a complete explanation of the relevant phenomena. Rather, what makes the problems easy is that, even though the solutions to these problems probably still require decades or even centuries of difficult empirical investigation, we nonetheless have every reason to believe that we can reach them using the standard methods of cognitive science and neuroscience. (Chalmers 1995, 1996) Solving the problem of attention, for example, simply awaits the empirical identification of a relevant neural mechanism. But what kind of mechanism could account for qualia? Though we strongly suspect that the physical system of the brain gives rise to qualia, we do not have any understanding of how it does so. The problem of accounting for qualia has thus become known, following Chalmers, as the hard problem of consciousness.

The hard problem of consciousness relates quite closely to what Joseph Levine had previously referred to as the explanatory gap. Given the scientific identification of heat with the motion of molecules, there is no further explanation that needs to be given: "our knowledge of chemistry and physics makes intelligible how it is that something like the motion of molecules could play the causal role we associate with heat.... Once we understand how this causal role is carried out there is nothing more we need to understand." (Levine 1983) In contrast, when we are told that pain is to be identified with some neural or functional state, while we have learned quite a bit, there is still something left unexplained. Suppose, for example, that we precisely identify the neural mechanism that accounts for pain—C-fiber firing, let's say. Still, a further question would remain: Why does our experience of pain feel the way that it does? Why does C-fiber firing feel like this, rather than like that, or rather than nothing at all? Identifying pain with C-fiber firing fails to provide us with a complete explanation along the lines of the identification of heat with the motion of...
molecules.

Some philosophers have claimed that closing the explanatory gap and fully accounting for qualia is not merely hard but rather impossible. This position, often referred to as *new mysterianism*, is most closely associated with Colin McGinn. According to McGinn, we will in principle never be able to resolve the mystery of what it is about the brain that accounts for qualia. (McGinn 1989) A similar, though slightly weaker, view is held by Thomas Nagel. According to Nagel, we currently do not have the conceptual apparatus necessary to even begin to understand how physicalism might be true. In order to solve the hard problem of consciousness, we would have to undergo a complete overhaul of our entire conceptual apparatus—a conceptual revolution so radical that we cannot even begin to conceive what the resulting *concepts* would be like. (Nagel 1998) But other philosophers reject the pessimism of the new mysterians as unwarranted or premature. Chalmers, for example, suggests that an explanation of how consciousness relates to the physical, even if it does not reduce to it, may well be enlightening. (See Chalmers 1996, 379)

It is perhaps easiest to see why the hard problem of consciousness is so hard by looking at particular attempts to account for qualia. The following three sections review three different theories of mental states—functionalism, physicalism, and representationalism—and the problems they face in accounting for qualia.

### 2. Qualia and Functionalism

The contemporary debate about qualia was framed in large part by discussions of functionalism in the late 1960s and early 1970s. Some attention had earlier been paid to qualia in connection with type identity theory, the view that mental state types could be identified with physical state types (for example, the mental state type pain might be identified to the neural state type C-fiber firing). But it was with the emergence of functionalism as a theory of mind that the debate about qualia began to heat up.

The intuition underlying the functionalist view is that the function of a mental state is its defining feature. Mental states are defined in terms of the causal role that they play in the entire system of the mind—that is, in terms of their causal relations to sensory stimuli, behavioral outputs, and other mental states. By defining mental states in this way, functionalism avoids many of the objections aimed at philosophical behaviorism, an early 20th century theory of mental states that defines them simply in terms of their input-output relations. Moreover, because a causal role can be defined independently of its physical realization (that is, because functional states are multiply realizable), functionalism avoids many of the objections aimed at the type identity theory. Rather than define pain in terms of C-fiber firing, functionalism defines pain in terms of the causal role it plays in our mental life: causing avoidance behavior, warning us of danger, etc., in response to certain environmental stimuli.

As plausible as functionalism may seem, however, it has long faced the charge that it is unable to account adequately for qualia. The causal role of a state seems to come apart from its qualitative aspects. To show this, opponents of functionalism have mounted two different kinds of arguments: (1) those aiming to show that two systems might be functionally identical even though only one of them has any qualia at all, and (2) those aiming to show that two systems might be functionally identical even though they have vastly different qualia from one another.

Falling in the first of these two categories, the *absent qualia argument* tries to establish that a system could instantiate the functional state of, say, pain without having any pain qualia. Ned Block originated this objection to functionalism with the thought experiment of the homunculi-headed robot (Block 1978).
Suppose a billion people were recruited to take part in a giant experiment. Each individual is given a very small task to perform—for example, to press a certain button when a certain light comes on. In doing so, each of them plays the causal role of an individual neuron, with the communications between them mirroring the synaptic connections among the neurons. Now suppose that signals from this network of people are appropriately connected to a robot body, so that the signals from the network cause the robot to move, talk, etc. If the network were set up in the right way, then it seems in principle possible that it could be functionally equivalent to a human brain. However, intuitively speaking, it seems very odd to attribute qualia to the robot. Though it might be in a state functionally equivalent to the state you are in when you have a pain in your right toe, it seems implausible to suppose that the robot is feeling pain. In fact, it seems implausible to suppose that the robot could have any phenomenal experience whatsoever. Thus, if the absent qualia objection is right, we can have functional equivalence without qualitative equivalence, so qualia escape functional explanation.

A related objection, falling into the second category, is the inverted qualia argument against functionalism, which arises from considering a possibility originally suggested by John Locke. Suppose that two people, Norma and Abby, are qualitatively inverted with respect to one another. Both of them refer to stop signs, Coke cans, and Elmo as "red," and both refer to sugar snap peas, Heineken bottles, and Kermit the Frog as "green." But Abby's phenomenal experience when she sees a Coke can is like Norma's phenomenal experience when she sees a Heineken bottle. When Norma sees the Coke can, she has a reddish experience; when Abby sees the Coke can, she has a greenish experience. Likewise, when Norma sees the Heineken bottle, she has a greenish experience; but when Abby sees the Heineken bottle, she has a reddish experience. Qualitatively, the two are inverted relative to each other.

Though most people find this scenario conceptually coherent, the functionalist can make no sense of this inversion. Abby and Norma both refer to the Coke can as "red." They both indicate that it is the same color as stop signs and ripe tomatoes. Functionally speaking, there is nothing to differentiate the states that Abby and Norma are in when they see the Coke can. But, by hypothesis, they have different qualitative experiences when they see the Coke can. Thus, it looks as if functional definitions of mental states leave out the qualitative aspects of mental states.

In response to these qualia-related objections, the functionalist might try to argue that we have not really imagined the scenarios that we think we have imagined. For example, can we really imagine what would happen if we had a billion people participating in a network to operate the robot? (In fact, even a billion people would not be enough to simulate the human brain, which is estimated to have 100 billion neurons.) Along these lines, William Lycan (1995, 50-52) argues that our intuition that the robot does not have qualia stems from a misguided focus on each microscopic part of the system rather than on the macroscopic system as a whole. Likewise, the functionalist might offer considerations to show that, contrary to how it first seems, the notion of behaviorally undetectable qualia inversion is not conceptually coherent after all. For example, because saturated yellow is brighter than saturated blue, the inversion between Norma and Abby would be detectable if they were both shown patches of saturated blue and saturated yellow and asked which was brighter. (See Tye 1995, 203-4)

Alternatively, if the functionalist cannot convince us that the absent qualia and inverted qualia scenarios are incoherent, he might instead narrow the scope of the theory, restricting it to mental states that are not qualitative. As John Haugeland argues, we can "segregate" the states that can be functionalized from the states that cannot: "if felt qualities are fundamentally different, so be it; explaining them is somebody else's business." (Haugeland 1978, 222) However, while this kind of segregation might save functionalism as a theory of cognition, it does so only by ignoring the hard problem of consciousness.
3. Qualia and Physicalism

As described above, the absent qualia objection and the inverted qualia objection specifically target functionalism, but they can be generalized to apply to physicalism more broadly. For the inverted qualia argument, the generalization is straightforward. Just as we can conceive of Abby and Norma being in functionally identical states, it does not seem implausible to suppose that their brains might be physically identical to one another. If so, then just as qualia escape functional explanation, they also escape physical explanation.

The generalization is less straightforward with the absent qualia argument. The homunculi-headed robot, though functionally identical to a human being, is not physically identical to a human being. However, in recent work, Chalmers has argued that we can conceive of what he terms "zombies"—beings who are molecule-for-molecule identical with phenomenally conscious beings but who are not themselves phenomenally conscious. In appearance and action, a conscious being and his zombie replica would be indistinguishable, but for the zombie, as Chalmers says, "all is dark inside." (Chalmers 1996, 96) When Zack and Zombie Zack each take a bite of chocolate cake, they each have the same reaction—they smile, exclaim how good it is, lick their lips, and reach for another forkful. But whereas Zack, a phenomenally conscious being, is having a distinctive (and delightful) qualitative experience while tasting the chocolate cake, Zombie Zack is experiencing nothing at all. This suggests that Zack's consciousness is a further fact about him, over and above all the physical facts about him (since all those physical facts are true of Zombie Zack as well). Consciousness, that is, must be nonphysical.

Chalmers' argument has the standard form of a conceivability argument, moving from a claim about conceivability to a claim about metaphysical possibility. Though zombies are probably not physically possible—not possible in a world that has laws of nature like our world—the fact that they are conceivable is taken to show that there is a metaphysically possible world in which they could exist. This form of argument is not entirely uncontroversial (see, for example, Hill and McLaughlin 1999), and there is also considerable debate about whether Chalmers is right that zombies are conceivable (see, for example, Searle 1997). But if Chalmers is right about the conceivability of zombies, and if this conceivability implies their metaphysical possibility, then it would follow that physicalism is false.

An early and influential discussion of the general problem that qualia pose for physicalism can be found in Thomas Nagel's seminal paper, "What is it like to be a Bat?" (Nagel 1974). Although it might be that not all living creatures have phenomenal experiences, we can be pretty confident that bats do—after all, they are mammals who engage in fairly sophisticated behavior. In Nagel's words, there is something that it is like to be a bat. But the physiology of bats is radically different from the physiology of human beings, and the way they interact with the world is radically different from the way that we interact with the world. What we do via vision, they do via echolocation (sonar). We detect objects by sight; bats detect objects by sending out high-frequency signals and detecting the reflections from nearby objects. Because this way of perceiving the world is so different from our own, it seems that their perceptual experiences must be vastly different from our own—so different, in fact, that Nagel argues that it is unimaginable from our perspective. We, who are not bats, cannot know what it is like to be a bat. Qualia are inherently subjective, and as such, Nagel argues that they cannot be accommodated by physicalism: "Every subjective phenomenon is essentially connected with a single point of view, and it seems inevitable that an objective, physical theory will abandon that point of view." (Nagel 1974, 520)

Related worries about physicalism and qualia have been forcefully developed by Frank Jackson in his well-known thought experiment involving Mary, a brilliant color scientist who has spent her entire life in a black-
and-white room. (Jackson 1982) Although she has normal color vision, her confinement has prevented her from ever having any color sensations. While in the room, Mary has studied color science through black and white textbooks, television, etc. And in that way she has learned the complete physical story about color experience, including all the physical facts about the brain and its visual system. She knows all the physical facts about color. But she has never seen anything in color. Now suppose that Mary is one day released from her room and presented with a ripe tomato. What should we imagine happens? Most people have the very strong intuition that Mary learns something from this perceptual experience. "Aha!" she might say. "Now I finally know what the color red is like."

The Mary case is the centerpiece of Jackson's knowledge argument against physicalism. While in the room, Mary knew all the physical facts about color, including the color red. When she is released from the room, Mary learns something about the color red, namely, what seeing red is like. What Mary learns consists of new, factual information. So there are facts about color in addition to all the physical facts about color (since Mary already knew all the physical facts about color). Thus, the argument goes, physicalism is false.

In the quarter century since Jackson's development of the knowledge argument, a vast literature has developed in response to it. Attempting to save physicalism, some philosophers deny that Mary learns anything at all when she leaves the room. If we really imagine that Mary has learned all the physical facts about color while in the room, then there would be no "Aha!" moment when she is shown a ripe tomato. We are led to think otherwise only because we typically fall short of imagining what we've been asked to imagine—we imagine only that Mary knows an immense amount about colors, that she has mastered all the information contained in our present science of color, which still remains incomplete. As Patricia Churchland has argued, "How can I assess what Mary will know and understand if she knows everything there is to know about the brain? Everything is a lot, and it means, in all likelihood, that Mary has a radically different and deeper understanding of the brain than anything barely conceivable in our wildest flights of fancy." (P.S. Churchland 1986, 332; see also Dennett 1991, 399-400)

Despite these reservations about what happens when Mary leaves the room, most philosophers—even most physicalists—accept Jackson's assessment that Mary learns something from her experience with the ripe tomato. Physicalists who grant this point have typically attempted two different strategies to respond to the knowledge argument: (1) They might accept that Mary gains new knowledge that isn't understood in terms of facts; or (2) they might accept that Mary's knowledge is factual but deny that she's learned anything new; rather, facts that she already knew are presented to her in a new way.

To pursue strategy (1), the physicalist might argue that the knowledge Mary gains when she leaves the room consists in nonfactual knowledge. Along these lines, David Lewis (1988) offers the ability hypothesis: When Mary leaves the room, all that happens is that she gains some new abilities regarding color that she didn't have before. Unlike before, Mary is now able to imagine, recognize, and remember the color red. So she gains know-how, but she doesn't learn any facts. Pursuing strategy (1) in a different way, Earl Conee (1994) offers the acquaintance hypothesis: When Mary leaves the room, all that happens is that she becomes acquainted with the color red. When you meet someone for the first time that you've previously heard or read a lot about, you don't necessarily learn any facts about them; rather, you just become acquainted with them. Conee thus argues that acquaintance knowledge (like ability knowledge) should not be understood in terms of facts. If either the ability hypothesis or the acquaintance hypothesis is right, and Mary does not learn any facts when she leaves the room, then the knowledge argument does not show that the physical facts are incomplete.

To pursue strategy (2), the physicalist might argue that Mary doesn't gain any new knowledge when she
leaves the room; rather, she simply comes to apprehend an old fact under a new guise. While in the room, she did not have the conceptual apparatus she needed in order to apprehend certain color facts in a phenomenal way. Having seen color, she has now gained new concepts—phenomenal concepts—and thus is able to re-apprehend the same facts she already knew in a different way. (Loar 1990) Whether there are genuinely phenomenal concepts, and if so, whether they do the work in answering the knowledge argument that the physicalists want them to, has recently been generating a growing literature of its own.

4. Qualia and Representationalism

While functionalism and physicalism are put forth as general theories of mind, representationalism aims specifically to give an account of qualia. According to this view, the qualitative character of our phenomenal mental states depends on the intentional content of such states. Representationalist views divide into two categories depending on exactly how they characterize this dependence. Weak representationalism makes a claim only about supervenience: The qualitative character of our mental states supervenes on the intentional content of those states (that is, if two experiences are alike representationally, then they are alike phenomenally). Strong (or pure) representationalism makes a further claim: The qualitative character of our mental states consists in the intentional content of such states. Strong representationalism thus offers a theory of qualia—it attempts to explain what qualitative character is. This section addresses the strong representationalist theory of qualia; hereafter, the modifier "strong" will be omitted.

Recall the distinction above between the easy problems of consciousness and the hard problem. Accounting for representational content is supposed to be one of the easy problems. It may take us an enormous amount of empirical work to get to the solution, but the standard methods of cognitive science will be able to apply. Thus, if qualia can be reduced to intentionality, then we have turned the hard problem of consciousness into an easy problem. A full and satisfactory account of qualia awaits only a solution to the easy problem of intentionality.

Consider pain qualia. Traditionally, philosophers classified pain experiences as non-intentional. However, the representationalist claims that this is a mistake. When one has a pain in one's leg, the experience represents damage in the leg. Moreover, its phenomenal feel—its painfulness—consists in its doing so. As Michael Tye argues, "[T]he phenomenal character of my pain intuitively is something that is given to me via introspection of what I experience in having the pain. But what I experience is what my experience represents. So, phenomenal character is representational." (Tye 1990, 338)

Given that the representationalist typically does not want to claim that all intentional content is qualitative, he must explain what is special about the intentional content in which phenomenal character is supposed to consist. My belief that Thomas the Tank Engine is blue and my mental image of Thomas the Tank Engine have similar intentional content; they both represent him as blue. So, what about the intentional content of the latter gives it its distinctive phenomenology? Here Tye has a particularly well-developed answer. He suggests that phenomenal content is a species of nonconceptual intentional content, in particular, nonconceptual intentional content that is poised and abstract. (Tye 1995) Because we can experience many things for which we lack concepts—for example, a proud parent might visually experience his young child’s drawing without having a concept for the shape that the drawing is—it is important that phenomenal content be restricted to nonconceptual content. The requirement that the contents be poised means that they "stand ready and in position to make a direct impact on the belief/desire system." (Tye 1995, 138) The requirement that the contents be abstract means that no particular concrete object is a part of them.

In support of their theory, representationalists often invoke what we might call the transparency thesis.
According to this thesis, experience is alleged to be transparent in the sense that we "see" right through it to the object of that experience, analogously to the way that we see through a pane of glass to whatever is on the other side of it. Gilbert Harman introduced such considerations into the contemporary debate about qualia in a now-famous passage: "When Eloise sees a tree before her, the colors she experiences are all experienced as features of the tree and its surroundings. None of them are experienced as intrinsic features of her experience. Nor does she experience any features of anything as intrinsic features of her experiences." (Harman 1990, 667) As Harman went on to argue, the same is true for all of us: When we look at a tree and then introspect our visual experience, all we can find to attend to are features of the presented tree. Our experience is thus transparent; when we attend to it, we can do so only by attending to what the experience represents. Representationalists contend that their theory offers the best and simplest possible explanation of this phenomenon. The best explanation of the fact that we cannot introspectively find any intrinsic features of our experience is that there are none to find; the phenomenal character of experience is wholly constituted by the representational content of the experience. (see especially Tye 1995, 2000)

Whether experience is really transparent in the way that the representationalists suppose has lately been the subject of some dispute, and there has also been considerable discussion about the relationship between experiential transparency and representationalism (See, for example, Kind 2003, Siewert 2004). Most problematic for the representationalists, however, has been the fact that their view falls victim to several persistent and compelling counterexamples. Many phenomenal states simply do not seem to be doing any representing—or, more cautiously, it seems that their phenomenal content far outruns their representational content. Ned Block has argued this point using the example of the orgasm: "Orgasm is phenomenally impressive and there is nothing very impressive about the representational content that there is an orgasm." (Block 2003, 543) He also discusses phosphene experiences, the color sensations created by pressure on the eyeball when one's eyelids are closed. Phosphene experiences do not seem to be representing anything; we don't take the experience to suggest that there are colored moving expanses out there somewhere.

Consider also the experience of seeing something flying overhead and hearing something flying overhead. While these two experiences have quite different phenomenal characters, their representational contents are plausibly the same: there's something flying overhead. (The most obvious way of differentiating them—by talking of the "way" of representing—brings in something nonrepresentational.) If this is right, then phenomenal character does not supervene on representational character. In response to objections of this sort, intramodal representationalists restrict their view so that it applies only within a given sensory modality. Unlike intermodal representationalists, who claim that all phenomenal differences, even differences between sensory modalities, can be explained in terms of representational content, intramodal representationalists think that we must offer some additional explanation to account for what makes a phenomenal experience auditory rather than visual, or visual rather than tactile. Typically, this additional explanation is provided in functionalist terms. (See Lycan 1996, esp. 134-35)

Along with these sorts of counterexamples, representationalism also falls victim to a version of the inverted qualia argument: the case of Inverted Earth (Block 1990). On Inverted Earth, the colors of objects are inverted relative to earth. Ripe tomatoes are green; unripe tomatoes are red. Big Bird is blue; the Cookie Monster is yellow. Other than this color inversion, everything else on Inverted Earth is exactly like earth. Now imagine that, without your knowledge, you are fitted with color-inverting lenses and transported to Inverted Earth. Since the lens-inversion cancels out the inversion of colors of Inverted Earth, you are unable to detect that you're in a different environment. When you look at the sky on Inverted Earth, you
have a blue experience even though the sky there is yellow; when you look at the green ripe tomatoes, you have a red experience. While originally on earth, your red experience while looking at ripe tomatoes represented red. But according to Block, after enough time passes and you have become embedded in the linguistic and physical environment of Inverted Earth, your reddish experience while looking at ripe tomatoes represents green (since that is the color of the ripe tomatoes on Inverted Earth). If Block's description of the Inverted Earth case is correct, then two experiences having identical qualitative character can differ in their intentional contents; thus, qualia do not supervene on intentional content and representationalism must be false.

In response to the Inverted Earth scenario, representationalists often adopt a teleological account of intentionality according to which the intentional contents of an individual's qualitative states are determined by the evolutionary history of its species. This allows them to reject Block's assertion that your intentional contents switch to match the Inverted Earthlings intentional contents. Humans have evolved such that red experiences represent red things. Thus, no matter how long you spend on Inverted Earth, the intentional contents of your reddish experiences will never switch to match the intentional contents of the Inverted Earthlings.

A completely different source of worry about representationalism has been raised by John Searle. Searle agrees with the representationalist that there is a close connection between phenomenal consciousness and intentionality, but he thinks that the representationalist gets the explanatory connection backwards. Rather than explain consciousness in terms of intentionality, Searle claims that we need to explain intentionality in terms of consciousness: "There is a conceptual connection between consciousness and intentionality that has the consequence that a complete theory of intentionality requires an account of consciousness." (Searle 1992, 132) Recent work by George Graham, Terry Horgan, and John Tienson argues along similar lines. On their view, "the most fundamental, nonderivative sort of intentionality is fully constituted by phenomenology." (Graham and Horgan 2008, 92; see also Horgan and Tienson 2002)

5. Eliminativism about Qualia

Rather than trying to find some way to fit qualia into a physicalist theory of mind, some philosophers have taken an entirely different attitude towards qualia. They deny that qualia exist. This position is known as eliminativism about qualia, and it commonly constitutes a part of a larger eliminativist project about mental states in general. For example, Paul and Patricia Churchland have argued (both together and individually) that as we gain more and more neuroscientific understanding of our mental lives, we will come to see that our current mental state concepts—belief, pain, sensation, qualia, etc.—all need to be discarded.

The Churchlands offer numerous useful analogies to help make this point. To consider just one of their examples: Ptolemaic theory placed the Earth at the center of the universe, around which a giant celestial sphere revolved. This created all sorts of difficult problems in need of solutions, like determining the cause of the sphere's rotation. When Newtonian theory displaced Ptolemaic theory, the notion of the celestial sphere was completely discarded. It wasn't that Ptolemaic theorists had an inadequate account of the celestial sphere; rather, what was discovered was that there was no celestial sphere. Thus, the problem of what causes the sphere's movement turned out to be a pseudo-problem. Similarly, the Churchlands predict that as our neuroscientific knowledge increases, we will come to see that the problem of qualia is a pseudo-problem, because we will come to see that there are no qualia—at least not as presently understood. Just as the celestial sphere did not turn out to be identifiable with or reducible to some element of Newtonian theory, qualia will not turn out to be identifiable with or reducible to some element of future neuroscientific
theory. Rather, the concept will have to be eliminated entirely. (P.S. Churchland 1986, 292-293; P.M. Churchland 1984, 43-45)

Insofar as eliminative materialism merely makes a prediction about what will happen once we increase our neuroscientific knowledge, it is hard to evaluate. However, Daniel Dennett offers related arguments for eliminativism designed to show there is such internal inconsistency in our notion of qualia that we are hopelessly misguided in trying to retain it. According to Dennett, there are no properties that meet the standard conception of qualia (that is, properties of experience that are intrinsic, ineffable, directly and/or immediately introspectible, and private). He reaches this conclusion by consideration of numerous thought experiments that are designed to tease out the alleged confusions inherent in our concept of qualia. For example, consider two coffee drinkers, Chase and Sanborn. Both discover one day that they no longer like the Maxwell House coffee they've long enjoyed. Chase claims: "Even though the coffee still tastes the same to me, I now no longer like that taste." In contrast, Sanborn claims: "The coffee now tastes different to me, and I don't like the new taste." But, asks Dennett, how do they know this? Perhaps Chase's taste receptors have changed so gradually that he hasn't noticed a change in taste; that is, perhaps he's really in the situation that Sanborn purports to be in. Or perhaps Sanborn's standards have changed so gradually that he hasn't noticed that he now employs different criteria in evaluating the coffee; that is, perhaps he's really in the situation that Chase purports to be in. There seems no first-person way for Chase and Sanborn to settle the matter, calling into question the idea that they have any kind of direct or special access to private properties of their experience. We might try to devise some behavioral tests to detect the difference, but if we could do so, that would suggest that qualia could be defined relationally, in reference to behavior, and this would call into question the idea that they are intrinsic. Thus, concludes Dennett, our conception of qualia is so confused that it would be "tactically obtuse" to try to salvage the notion; rather, we should just admit that "there simply are no qualia at all." (Dennett 1988)

6. Naturalistic Dualism

There is at least one further option available to philosophers when confronting the hard problem of consciousness. Without denying the reality of qualia, one might simply accept that they resist reduction in physical, functional, or representational terms and embrace some form of dualism. This is David Chalmers' own approach to the hard problem. Because he believes that we can account for phenomenal consciousness within a solely natural framework, he adopts what he refers to as naturalistic dualism.

Descartes' dualism was a version of substance dualism. According to Descartes, the mind is an immaterial substance existing independently of the body. In contrast, Chalmers' dualism is a version of property dualism. This view does not posit the existence of any nonphysical or immaterial substances, but instead posits the existence of properties—qualia—that are ontologically independent of any physical properties. Though these properties are not entailed by physicalism (that is, though they do not logically supervene on physical properties) they may nonetheless somehow arise from them. As Chalmers describes his view: "[C]onsciousness arises from a physical substrate in virtue of certain contingent laws of nature, which are not themselves implied by physical laws." (Chalmers 1996, 125)

Physics postulates a number of fundamental features of the world: mass, spin, charge, etc. Naturalistic dualism adds nonphysical phenomenal properties to this list. Correspondingly, it suggests we must add fundamental laws governing the behavior of the fundamental phenomenal features to the list of the fundamental laws governing the behavior of the fundamental physical features of the world. We don't presently understand exactly what these new laws and the completed theory containing them will look like,
and Chalmers admits that developing such a theory will not be easy, but in principle it should be possible to do so.

This commitment to lawfulness is what allows Chalmers to remain within a naturalistic framework, even as he abandons the physicalistic framework. On his view, "the world still consists in a network of fundamental properties related by basic laws, and everything is to be ultimately explained in those terms. All that has happened is that the inventory of properties and laws has been expanded [beyond the physical properties and laws]." (Chalmers 1996, 127-8) In a similar spirit, Gregg Rosenberg has recently offered a view he calls liberal naturalism. Though liberal naturalism holds that the fundamental properties of the world "are mutually related in a coherent and natural way by a single set of fundamental laws," it denies that these properties and laws can all be completely captured in physical terms. (Rosenberg 2004, 9)

In giving up physicalism, naturalists argue that we can retain almost everything that's important about our current scientific worldview. But the adoption of nonphysicalistic naturalism typically leads in two directions that many have thought problematic. First, it seems to imply panpsychism, the view that everything in the universe has consciousness. Once you accept the existence of nonphysical features of the world that are fundamental, it is hard to find a principled way of limiting exactly where those fundamental features are found. As Chalmers admits, "if experience is truly a fundamental property, it seems natural for it to be widespread." (Chalmers 1996, 297; see also Nagel 1979) Second, it seems to commit one to epiphenomenalism, the view that qualia lack any causal power whatsoever. Intuitively, we believe that the qualitative character of pain—the fact that it hurts—causes us to react the way that we do when we feel pain. But if qualia are epiphenomenal, then the painfulness of pain is causally inert.

In addressing the first of these two worries, Chalmers denies that naturalistic dualism entails panpsychism. Though he recognizes that it provides a particularly elegant way of working out the details of the view that experience supervenes naturally on the physical, he believes that there remains the possibility that those details could be worked out another way. Benjamin Libet, for example, offers a theory that sees consciousness as fundamental without endorsing panpsychism (Libet 1996). In contrast to Chalmers and Libet, Rosenberg concedes that nonreductive naturalism will most likely require us to adopt at least a weak form of panpsychism, and he offers arguments to show why this consequence should not be seen as threatening.

Even if naturalism leads only to a mild form of panpsychism, however, most contemporary philosophers would find this extremely problematic. How could blades of grass, or rocks, or atoms be conscious? Panpsychism is almost universally regarded with skepticism, if not outright scorn. Colin McGinn, for example, has claimed that panpsychism is "metaphysically and scientifically outrageous." (McGinn 1996, 34) Similarly, in reaction to Chalmers' panpsychist musings, John Searle calls panpsychism "absurd" and claims that there is "not the slightest reason" to adopt it. (Searle 1997, 161)

The worries about epiphenomenalism are no less troublesome for the naturalist than are the worries about panpsychism. Intuitively speaking, qualia are important aspects of our mental lives. The itchiness of an itch makes us scratch, the delicious taste of chocolate leads us to reach for another piece, the wrenching feeling of grief erupts in a flood of tears. But if qualia are physically irreducible, then it seems they must be left out of the causal explanations of our actions. We typically assume that the physical world is causally closed; all physical events, including bodily movements, can be given complete causal explanations in wholly physical terms. Unless we reject causal closure, then assuming we do not want to embrace the possibility of causal overdetermination, qualia have no role to play in the causal story of our actions.
We can easily see why naturalism leads to epiphenomenalism by reconsidering the zombie world. By hypothesis, your zombie twin is behaviorally indistinguishable from you despite having no qualia. His actions can be causally explained entirely by the physical workings of his brain. But he’s a molecule-for-molecule duplicate of you, so the physical workings of your brain can provide a complete causal explanation of your actions. Your qualia play no role in causing the actions that you perform.

Chalmers addresses the threat of epiphenomenalism in two ways. First, he suggests that our inadequate understanding of the nature of causation may here be leading us astray: "it is possible that when causation is better understood we will be in a position to understand a subtle way in which consciousness may be relevant." (Chalmers 1996, 150) Second, he tries to show that epiphenomenalism may not be as unpalatable as many have thought. In particular, he argues that we don’t have any reasons to reject epiphenomenalism except for its seeming counterintuitive; there are no effective arguments against it. (See also Jackson 1982.) Moreover, given the fatal flaws that threaten the competing alternatives to naturalistic dualism, it may turn out that accepting some degree of counterintuitiveness is the small price we have to pay in order to develop a coherent and unmysterious view of consciousness and its place in nature.

7. References and Further Reading


A very useful collection bringing together Block’s impressive body of work in philosophy of mind on issues relating to functionalism, qualia, and consciousness.


A helpful characterization of the issues surrounding representationalism (which Block calls representationism) and a defense of a qualia realist view he calls phenomenonism.


A reply to Harman's "The Intrinsic Quality of Experience." This paper introduces the much-discussed Inverted Earth thought experiment, a version of the inverted qualia argument targeting representationalism.


An influential work that develops in detail the absent qualia objection to functionalism.


An anthology collecting much of the classic work on consciousness.


A very useful overview of the issues surrounding representationalism.


One of the most important books in philosophy of mind over the last twenty years; introduces and discusses in detail the hard problem of consciousness. Although the book is technical in parts, the most technical sections are indicated by asterisk and can be skipped without losing the overall argument.


An accessible introductory text to the philosophy of mind, though Churchland's own eliminativist leanings shade his treatment of the issues discussed.


A classic presentation of Conee's "acquaintance hypothesis" in response to Jackson's knowledge argument.


Argues, in Dennett's characteristically jocular style, for eliminativism about qualia.


A sustained argument for representationalism, with sustained discussion of how representation works.


Introduces considerations of the transparency of experience into contemporary discussions of qualia.


Jackson's classic paper first laying out the Mary case and the knowledge argument against physicalism.


Reviews the history of the use of the term "qualia," both before and after C.I. Lewis introduced it into the philosophical literature in roughly its contemporary sense.


Introduces the term "qualia" in its contemporary sense (introspectible, monadic, subjective properties), though Lewis uses it in the context of sense data.

An influential presentation of the "ability hypothesis" in response to Jackson's knowledge argument.


An anthology that collects Jackson's original two papers laying out the knowledge argument along with many important papers in response. Also contains Jackson's recent surprising recantation of the original argument, published here for the first time. Jackson now believes that the representationalist view helps us to see how the argument goes wrong.


A development of Lycan's intramodal representationalism.


Defends new mysterianism, that is, the view that the problem of consciousness cannot in principle be solved.


A classic paper arguing that physicalism cannot accommodate the subjective aspects of experience—much-cited and well worth reading.


Argues for a conceptual connection between consciousness and intentionality.


An interesting argument attempting to show that functionalism can handle inverted qualia. Shoemaker's own view about qualia is somewhat idiosyncratic in that he denies they are directly introspectible.


Further development of the representationalist view, including responses to common criticisms of the view.


Develops a strong representationalist view in an attempt to unravel several puzzling aspects of consciousness (its subjectivity, transparency, etc.).


An early statement of representationalism, here limited specifically to pain.
Author Information

Amy Kind
Email: amy.kind@cmc.edu
Claremont McKenna College
U. S. A.

Copyright © The Internet Encyclopedia of Philosophy. All rights reserved.