

## CHAPTER 2: LEVINE AND BLOCK

### 2.1 Levine's Explanatory Gap

#### 2.1.1 Levine's Defense of Materialism

Joseph Levine rejects Chalmers's model of reductive explanation and his antimaterialist metaphysical conclusion. However, he argues for an epistemological problem of consciousness, a problem exemplified by the apparent failure of any convincing materialist explanation. In his 2001 book Purple Haze: The Puzzle of Consciousness and in a number of related papers (1983, 1993), Levine details both his support for materialism and his worry over an "explanatory gap" between materialist theory and our first-person understanding of conscious experience. So if Levine is correct, even if Chalmers's metaphysical hard problem has a solution, there is still a special explanatory problem when it comes to consciousness. In this section, I will lay out Levine's materialist response to Chalmers, and then I will present his formulation of an epistemic problem of consciousness. I will conclude by criticizing Levine's claim and presenting the residual challenge posed by his view.

Levine begins his defense against Chalmers by offering the following characterization of materialism:

Only non-mental properties are instantiated in a basic way; all mental properties are instantiated by being realized by the instantiation of other, non-mental properties (2001, 21).

This characterization is crafted to avoid a worry about formulating materialism in terms of our current physics.<sup>1</sup> It holds that whatever the basic properties turn to be, if materialism is true, they won't include mental properties. However, if Chalmers's arguments are correct, then phenomenal consciousness fails to supervene on the physical and must be counted among the basic features of reality. Thus, materialism as characterized by Levine would still be false.

However, Levine contends that we have positive reason to believe in materialism, even in the face of the apparent difficulties. Materialism, according to Levine, is the best explanation of the causal efficacy of the mental. He argues as follows. An extremely fruitful assumption of modern physics holds that the physical world is "causally closed"--that is, every physical effect is fully accounted for by a physical cause. But on occasion it seems obvious that the mind causes the body to move. But to avoid violating causal closure, the mind too must be physical. Therefore materialism must be true. The alternatives are dualism and epiphenomenalism. Dualism rejects causal closure, but it is saddled with explaining how non-physical substance or properties causally interact with a physical ones. Epiphenomenalism respects causal closure by holding that the mind is not causally efficacious. Levine argues that dualism's interaction problem is insurmountable, and that epiphenomenalism is highly counterintuitive. Therefore, he concludes that materialism is justified as the best explanation of mental causation.

However, Chalmers's arguments against materialism still must be defused. As noted in chapter 1, one of Chalmers's central arguments against

---

<sup>1</sup> See Montero 1999.

materialism invokes zombies, holding that the mere conceivability of zombies implies their possibility, undermining materialism. Levine focuses his attack on this claim. On Chalmers's model, zombies are conceivable because they are not ruled out a priori by our concepts. Further, our concepts, according to Chalmers, define the space of logical possibility--flying telephones are possible, male vixens are not; this is a catalog of what is logically possible. Zombies are conceivable, so they are possible. But this entails the failure of logical supervenience--there is a possible world with the same physical facts as our world, but different mental facts--falsifying materialism.

Levine challenges the theory of conceptual content underwriting the link between conceivability and possibility in Chalmers's argument. According to Chalmers, our concepts are (in part) constituted by a priori accessible semantic connections.<sup>2</sup> It is an a priori accessible feature of our concept "water" that water is a liquid, for example. This underwrites our ability to evaluate would-be supervenience claims. For instance, because we know, by knowing the intension of the term, that water is the clear, potable liquid that fills our rivers and lakes, we can ascertain a priori the truth of the supervenience conditional "if H<sub>2</sub>O is the clear, potable liquid that fills our rivers and lakes, then water is H<sub>2</sub>O." Further, it is not conceivable, once we discover that H<sub>2</sub>O plays the "watery role," that there could be H<sub>2</sub>O present but no water present. The intension of "water" just is the "watery role," and if we have the role-filler, it is inconceivable that we don't have water, according to Chalmers. A priori accessible semantic information is crucial

---

<sup>2</sup> "In part" because Chalmers's holds that the secondary intension of a concept need not be a priori accessible. Primary intensions are what matter for Chalmers's antimaterialist arguments. See Chalmers 1996, chapter 2, 1999.

to Chalmers's model. Levine argues that such information is not available, and so the model fails.

Levine argues for a rival account of conceptual content. On his view, concept possession requires standing in the right causal or nomological relation to the concept's referent. It does not require possessing a rich network of semantic information about the concept's referent. On this view, to have the concept "water" is to have a representational state that is appropriately caused by water or that is lawfully connected to the presence of water. Thus, one can possess the concept "water" without knowing that water is a liquid, is potable, is clear, etc. While we may have the belief that water is a liquid, this is not constitutive of concept possession, nor is it information that is a priori accessible on the basis of concept possession alone. We require empirical investigation to establish that water is a liquid. This has serious consequences for Chalmers's arguments.

Chalmers's position is premised on the fact that we can distinguish between possible and impossible situations exclusively on the basis of conceivability intuitions. Once we have analyzed "water" and discovered that H<sub>2</sub>O fills that role, it is inconceivable that a sample of H<sub>2</sub>O is not a sample of water. But analyzing "consciousness" and positing that it is realized by a physical or functional property does not render zombies inconceivable. Therefore, Chalmers concludes, water logically supervenes on the physical, but consciousness does not. However, Levine argues that zombie-H<sub>2</sub>O (H<sub>2</sub>O that is nonetheless not water), like zombies, is in fact conceivable. Both Chalmers and

Levine agree that zombie-H<sub>2</sub>O is impossible. Thus, its conceivability would break the link between conceivability and possibility needed for Chalmers's model, thereby blocking his antimaterialist conclusion.

Levine's theory of content explains the conceivability of zombie-H<sub>2</sub>O. Given that our mental terms "water" and "H<sub>2</sub>O" in fact refer to the same stuff, it is impossible that H<sub>2</sub>O is present and water is not. However, if all that matters for concept possession is causal/nomological relations, we cannot tell simply by a priori reflection that they co-refer. We must investigate the world to determine what causal/nomological connections hold. For all we know before that, "water" may refer to clowns and "H<sub>2</sub>O" to accordions. In this case, it's conceivable that H<sub>2</sub>O is present and water is not. But the scenario is not possible. Thus, the link between conceivability and possibility needed for Chalmers's argument is broken, and materialism is saved.

Levine's argument in effect expands the space of conceivable situations to include all situations that are not contradictory on the basis of logical form alone. In essence, this is a rejection of the a/s distinction, though one rooted in considerations from causal theories of content rather than confirmation holism. But the result is largely the same. Chalmers's model cannot rule out a wide variety of conceivable situations that are nonetheless deemed impossible by all parties involved. Conceivability is not a reliable guide to possibility, so the mere conceivability of zombies (and the inverted spectrum) does not threaten materialism.<sup>3</sup>

---

<sup>3</sup> Levine's argument thus provides an additional attack on Chalmers's use of a priority and analyticity. However, Levine's position is committed to an atomistic view of content, while my

### 2.1.2 The Explanatory Gap

Despite his rejection of Chalmers's metaphysical antimaterialist conclusion, Levine argues that consciousness still poses a special epistemic problem for a materialist science of the mind. At the core of Levine's argument are the same conceivability intuitions used by Chalmers against materialism. Levine holds that the persistence of these intuitions, even in the face of viable empirical theory, points to a serious shortcoming with materialism. We have good reason to believe in the truth of materialism, but when presented with materialist theory, we still are left wondering how the physical brain could be phenomenally conscious. In other examples of scientific explanation, once a theory is developed, our anti-theoretical intuitions are blocked. But we can still legitimately ask why the brain is conscious, even if we accept materialism. Levine calls this problem the "explanatory gap" (1983, 1993, 2001). Even when presented with an apparently true theory of the physical brain, we are left wondering about consciousness. There is an epistemic gap between physical theory and the phenomenal mind.

Levine begins his argument for the presence of an explanatory gap by clarifying what he takes to be required of a good scientific explanation. He writes,

in a good scientific explanation, the explanans either entails the explanandum, or it entails a probability distribution over a range of alternatives, among which the explanandum resides... [W]e achieve

---

arguments relied on holistic considerations. Further discussion of these important issues concerning content goes beyond the scope of this dissertation.

understanding when we can see why, given the information cited in the explanans, the phenomenon cited in the explanandum had to be; or, to put it another way, why the relevant alternatives are ruled out, as inconsistent with the explanans (2001, 74).<sup>4</sup>

Thus, we ought to be able to deduce the explanandum in a successful explanation, according to Levine. He argues that if we cannot construct the desired deduction, there are three possible explanations. One is that we haven't fully specified the mechanisms and processes cited in the explanans. Two is that the target phenomenon is stochastic in nature, and the best that can be done is delivering a range of probabilities concerning the explanandum. The third is that there are as yet unknown factors at least partially involved in determining the phenomenon. If we've adequately specified the mechanisms and processes in question, and if we've adjusted for stochastic phenomena, then their description should deductively entail the explanandum, or the third possibility is in effect. But the third possibility is "precisely an admission that we don't have an adequate explanation" (2001, 76). Thus, if the explanandum is not entailed by the explanans, we don't have an adequate explanation.

---

<sup>4</sup> Levine argues that his view of explanation is not only compatible with the deductive-nomological model of Hempel (Hempel 1965), but that it also fits with Salmon's "ontic" conception of explanation (Salmon 1989). The ontic conception holds that in order to explain a phenomenon we must exhibit the mechanisms that are causally responsible for it. Levine contends that if we produce a description of the appropriate mechanism, we should be able to deduce a description of the phenomenon. That preserves the spirit of the ontic model and maintains the deductive requirement. Some argue against a deductive requirement because deductions are symmetric while explanations are not (Achinstein 1981); others argue that actual deductions are never forthcoming, and we always employ a variety of pragmatic considerations in successful explanation. I will say more about the requirements on an explanation of consciousness at the beginning of chapter 3. For now, I will accept Levine's claim; my explication and criticisms of his view are independent of this issue.

Levine argues that the conceivability of zombies, the conceivability of the inverted spectrum, and the problem of other minds (for entities whose constitution is very different from our own), all show that the deduction required for a successful explanation of consciousness is lacking. If we possessed the deduction, we could rule out zombies and inversion on that basis. For example, if we knew that conscious experiences of red were identical to c-fiber firings, and we knew that a creature had firing c-fibers, we would know that it was having a conscious experience of red. We could deduce the presence of conscious states in the creature, and this would rule out both zombies and inversions. Likewise, we would be in a position to judge if it had conscious states at all, solving the problem of other minds. However, despite the development of a number of reductive materialist theories and a great improvement in our knowledge of the brain, the problematic cases persist. Zombies and inversion seem as easily imagined as ever, and the problem of other minds still seems pressing. We cannot, it seems, deduce the presence of consciousness from a materialist description of the mind. Therefore, we have not explained consciousness.

Still, there is a reasonable explanation for the persistence of these intuitions. A materialist explanation of consciousness entails that certain brain states are identical to certain phenomenal states. The example above posited that red experiences are identical to C-fiber firings. Levine's questions amount to asking, how could this identity be true--what explains it? But ordinarily, identity claims do not have explanations. A thing just is what it is. Mark Twain, for example, is Samuel Clemens. There is no sense to the question, "how could

Mark Twain be Samuel Clemens?" He just is. We may ask why we thought that one thing was really two. But once that story is filled-in, the question does not make sense. So the explanatory story that Levine demands need not and could not be given. He is asking how something could be itself, and there is nothing meaningful to say about that.<sup>5</sup>

Levine counters that when we closely compare identities involving consciousness to other examples, we see that they still fail to eliminate meaningful questions. He canvases identity claims involving indexicals, demonstratives, and natural kinds, and concurs that in those cases the questions do indeed drop away. However, the fact that meaningful questions remain in the case of consciousness suggests a substantive difference, one indicative of the explanatory gap. In the case of indexicals, it is widely agreed (following Perry) that indexical claims cannot be derived from nonindexical descriptions. However, this need not point to anything of metaphysical interest, and indeed, when the relevant referents are introduced, questions fall away. Consider the claim "I am here now." Once we learn that Josh Weisberg is the current referent of "I," my office is the current referent of "here," and 9:45 is the current referent of "now," there isn't any sense in asking, "But how could I be here now?" given that Josh Weisberg is in my office at 9:45. That's just how these terms work, and once the relevant contextual features are supplied, meaningful questions drop away.

The situation is the same with demonstratives. We arguably can't derive their referents from nondemonstrative descriptions, but once the relevant

---

<sup>5</sup> This claim is pressed by Papineau 1995, 2002; Block, 2002a, 2002b; and others.

contextual features are filled-in, certain questions don't make sense. Following Levine's example, imagine that I point blindly in front of me and say, "I wonder what that is?" I look up and it's a red diskette case. Does it make any sense to ask how it could be that my red diskette case is that? Levine holds that it does not. Once we've determined what lies at the end of the demonstrative, this kind of question lacks sense.

More to the point is the case of natural kind identities, like "water = H<sub>2</sub>O." In arguing against Chalmers, Levine defended a causal/nomological theory of content. The view entails that prior to empirical investigation, concepts are simply uninterpreted strings, mere labels in the language of thought. We must discover how we are hooked to the world in order to determine the referents of our concepts. But once that information is supplied and we discover that two "mental words" refer to the same thing, it makes no sense to ask, "But how could they co-refer?" If the referential links are there, they just co-refer, and that's that. In the case of "water," we learn it refers to a particular stuff in our environment. Then we learn that a chemical term "H<sub>2</sub>O" refers to the same stuff. We may ask how it is that H<sub>2</sub>O displays the various features that we already believed were displayed by water. But once the requisite chemical theory is provided, those questions have good answers. We learn that both terms refer to the same stuff, and meaningful questions drop away.

These examples are in stark contrast to the case of consciousness, according to Levine. We can still meaningfully ask how C-fibers firings (or specific functional or representational states) could be my conscious experience

of red. We may fully believe that the two terms co-refer. But this does not eliminate the questions. Thus identity claims involving consciousness differ from ordinary cases of indexicals, demonstratives, and natural kinds.

Still, there is a crucial unresolved point. In other cases of natural kinds, prior to the development of the requisite background theory (chemistry, in the water example), meaningful questions of a sort were indeed still open. Before the development of chemical theory, we could reasonably ask why we should think that "water" and "H<sub>2</sub>O" refer to the same stuff. How could this molecular stuff be transparent, thirst-quenching, etc.? Granted, once the theory was in place, such questions fell away, but before that, there would appear to be a gap between our everyday knowledge of water and the claims of chemical theory. How do we know that this isn't the case with consciousness? Why think that the background theory is in place to block our questions, especially given our relatively scant knowledge of the brain?

Levine responds that there is an important difference in the way we access our conscious mental states and the way we access everything else. Given this difference, it is apparent that prospective reductive explanations fail to narrow the gap. Our access to the world by way of concepts like "water," "H<sub>2</sub>O," "cat," etc., involves what Levine calls "thin modes of presentation" (MOPs). Thin MOPs

merely label a phenomenon/substance in the world... [This] simultaneously explains why water facts are not strictly derivable from the physical facts and also why, nevertheless, requests to explain the identity

of water with H<sub>2</sub>O, once the relevant physical facts are known, are unintelligible. There isn't enough cognitive substance associated with 'water' to make sense of this request for explanation (2001, 84).

I termed these thin MOPs "uninterpreted strings in the language of thought" above. We cannot a priori derive them from the physical facts because we do not know a priori what they refer to. But once empirical information is supplied, there isn't any room for questions about how two terms could co-refer. If they stand in the right causal/nomological relationships, they co-refer.

On the other hand, our awareness of our conscious mental states is by way of "thick modes of presentation." With conscious experience, according to Levine, "We are not just labeling some 'we know not what' with the term 'reddish,' but rather we have a fairly determinate conception of what it is for an experience to be reddish" (2001, 84). Levine calls this sort of conception "substantive and determinate." He writes,

When I think of what it is to be reddish, the reddishness itself is somehow included in the thought; it's present to me. This is what I mean by saying it has a 'substantive' mode of presentation. In fact, it seems the right way to look at it is that reddishness itself is serving as its own mode of presentation. By saying the conception is 'determinate,' I mean that reddishness presents itself as a specific quality, identifiable in its own right, not merely by its relation to other qualities (2001, 8).

Because we possess this special access to our conscious mental states, materialist identity claims fail to close off meaningful questions. The proposed

identities do not answer to our substantive and determinate conception--indeed, we cannot see how they could. Levine calls an identity that leaves open meaningful questions "a gappy identity" (2001, 81ff). If the deduction from the physical facts to an explanatory target runs through a gappy identity, we are left with an explanatory gap. Such is the case with consciousness.

Levine's insistence on thick MOPs for conscious mental states underwrites his rejection of a range of materialist explanations of consciousness (2001, chapter 4). He argues that theories which construe conscious qualitative mental properties, or "qualia," as relational fail to do justice to our deep intuition that qualia are intrinsic features of experience. This intuition is underwritten by our substantive and determinate access to qualia--we have seemingly direct access to a feature which apparently defies materialist explanation. Furthermore, so-called "higher-order" theories of consciousness hold that our access to our conscious states is mediated by representation. But according to Levine, the substantive nature of thick MOPs demands that qualia are somehow constitutively involved in our awareness of them, blocking higher-order views. Finally, neuroscientific theories of qualia fall short as well. Neuroscience deals in relational information, for example data concerning neural connections and firings or the flow of information through neural nets. But relational information cannot do justice to our substantive and determinate awareness of the intrinsic features of experience, according to Levine. Since neuroscience arguably deals only in relational information, it cannot close the explanatory gap.

Thus it seems that there is a special explanatory problem of consciousness, even if we accept a materialist metaphysics. Our special access to consciousness undermines the possibility of a relational explanation, and it seems that relational information is all that the materialist has to offer. However, Levine's claim begs a crucial question. Why should we accept that our so-called substantive and determinate MOPs provide us with a fully accurate grasp of the nature of our mental states? Perhaps such states only appear to have intrinsic experiential character, a character that can't be fully explained in a materialist theory. Levine requires a defense of the claim that things are as they appear when we access our conscious states, and further, that there are no hidden workings that explain the nature of consciousness. Otherwise, it is open for his opponent to charge that despite appearances, conscious states really are functional or neurological states, albeit states that we do not recognize as such in introspection. Levine thus seems committed to the Cartesian doctrine that the mind is transparent to its subject, and the subject is authoritative, or even infallible, about the contents of her mind.<sup>6</sup> In keeping with the commonsense approach to fixing the data argued for in chapter 1, claims that go beyond the folk-psychological appearances to the underlying nature of consciousness require argument. It is not enough to say that conscious states appear to have an intrinsic quality, therefore they do.

---

<sup>6</sup> Levine acknowledges that his view commits him to "self intimating" conscious states, the idea that qualitative character is intrinsically conscious. But he does not argue for this, and holds that it is part of the "paradoxical nature" of the link between our subjective access to our conscious states and their character (2001, 109). He requires support for this claim, and for the stronger claim of infallibility.

However, Levine does not directly argue for such a claim. He comes closest in rejecting the idea that he, himself, might be a zombie, writing

it is not really even conceivable to me that I might be a zombie. I can rule out this possibility from within... First-person skepticism doesn't even get a foothold because my epistemic situation in some way includes my conscious experience... The qualia are essential components of how, cognitively and epistemically, it is with me (2001, 167).

However, this clearly does not establish transparent access--access where we can "see right through" to the nature of our conscious states--nor does it entail that we are infallible with respect to the underlying reality of mental states. It is still possible that the features we are aware of are relational in nature. Further, Levine explicitly rejects any appeal to privileged access and infallibility or incorrigibility (103, 117, 138ff). This is in keeping with the commonsense view of the mind--we can be in error about what states we are in, and we are not in infallible contact with the underlying nature of the mind. But it is not clear that he can do without such an appeal.

In its place, Levine offers the following concerning intuitions about the apparently intrinsic nature of qualia and our determinate and substantive access to them:

On my view, intuition has no special epistemic status; it's not a faculty in its own right, nor are its dictates to be treated as incorrigible. As far as I can see, intuition is just reasonableness. That is, to say that something is intuitively wrong or odd is to say that it strikes one as unreasonable,

implausible. One could be wrong about this, and the basis for this response should always be sought out to the degree possible, but sometimes one just has to rest on the fact that some hypothesis seems blatantly implausible (2001, 103).

But this certainly leaves open the possibility that, despite appearances, qualia are in fact relational, and, therefore, there is no explanatory gap. Our conscious states may appear to have intrinsic features, but that may just be appearances; the reality may be quite different. And our "substantive and determinate" access may only appear to constitutively involve the conscious states we access. Indeed, we may have no insight into the nature of the access. All we can reasonably conclude is that the access appears substantive and determinate and that we appear to access intrinsic features, while the reality may be quite different. In the absence of an argument against this possibility, it is an open question whether there really is an explanatory gap.

There is an additional move, however, that might seem to aid Levine. While folk psychology may not license infallible access to the states we are in, it may provide infallible access to how things seem to us in conscious experience. I may be incorrect that I am in some sensory or emotional state, but can I really be wrong that I seem to be in those states? And perhaps this "seeming" is itself an intrinsic property of my experience, thus establishing the presence of a gap-generating feature. But this move does not help. Even if we are infallible in this manner (and I will argue in later chapters that even this is unfounded), all that this provides Levine is that we seem to be in states with intrinsic properties. I

concur, and hold that this appearance requires explanation. But I resist the move from it seeming we are in states with intrinsic properties to the claim that we are in states with intrinsic properties. Why think the "seemings" themselves are intrinsic? While I may be infallible that I seem to be in pain, how does that establish that I am infallible (or even reliable) about whether the "seemings" themselves are intrinsic or relational? Nothing in folk psychology licenses this claim. Indeed, Armstrong's point seems fully appropriate here: just because we are not aware that the seemings are relational, this does not entail that we are aware of them as nonrelational. In the absence of additional support for this shift, we have no reason to accept it. There is nothing to be gained by Levine in this move of "introspective ascent."

To recap, there appears to be an explanatory gap because of the nature of our access to our conscious states. We seem to be aware, in a direct and unmediated manner, of a determinate quality "identifiable in its own right." If we take this access as fully accurate, materialist explanations involving functional or neurological states will inevitably seem to fall short. Such theories will have the task of explaining an intrinsic quality of experience in relational terms. But in the absence of supporting argument, we need not take this access as fully accurate concerning the nature of our conscious states. Perhaps they only appear to have an intrinsic quality, while in fact they do not. Without an argument supporting the accuracy claim, it is an open possibility that our conscious states may be explained in functional or neuroscientific terms. Levine provides no such

argument; therefore, I conclude that there may only appear to be an explanatory gap.

We are left with the following task. We must explain, in materialistically acceptable terms, the apparently substantive and determinate access we have to seemingly intrinsic features of conscious experience. This opens the door to a satisfying materialist theory of consciousness--one that fully accounts for the commonsense data while accounting for the intuitions that prompt claims of hard problems and explanatory gaps. I will have more to say about the requirements for such a theory of access at the beginning of chapter three. However, there is another theorist, Ned Block, whose work seems to point to a special explanatory problem of consciousness, one that differs from other scientific conundra. Further, he rejects both Chalmers's hard problem and Levine's explanatory gap. However, his defense of materialism incurs what he terms "the harder problem of consciousness." I will turn to his work in the next section.

## **2.2 Block's Phenomenal Consciousness**

### **2.2.1 Kinds of Consciousness**

In his 1995 paper "On a Confusion about a Function of Consciousness," Ned Block writes

The concept of consciousness is a hybrid or better, a mongrel concept: the word "consciousness" connotes a number of different concepts and denotes a number of different phenomena. We reason about "consciousness" using some premises that apply to one of the

phenomena that fall under "consciousness," other premises that apply to other "consciousnesses," and we end up with trouble (Block 1995, 375). He identifies four different phenomena that are referred to by "consciousness," most importantly one that defies characterization in "cognitive, intentional, or functional terms" (1995, 381). Confusing this kind of consciousness with one of the others leads to the false sense that the mystery of consciousness is easily solvable by the methods of a functionalist cognitive science or neuroscience, which deal in cognitive, intentional, and functional explanation. Thus, if we are to avoid illicitly explaining away the problems of consciousness, we must pay careful attention to Block's theoretical distinctions.

The first kind of consciousness Block identifies is "phenomenal consciousness" or "phenomenality" (p-consciousness, for short). P-consciousness is a "pretheoretic" notion characterized by mental states that there is something it is like for the subject to be in. P-consciousness is "what it is like to have an experience. When you enjoy the taste of wine, you are enjoying gustatory phenomenality" (Block 2001, 202). The paradigm cases of p-conscious states are sensations, states with a sensory or qualitative character. Block acknowledges that this is not a particularly informative way of characterizing the phenomenon, but that is to be expected. On his view, p-consciousness cannot be characterized in a noncircular way. We can only state our characterization of p-consciousness in terms of closely synonymous expressions. P-conscious states are thus "experiences" or "states that there is something it is like for the subject to be in." Furthermore, and most crucially, p-consciousness, according to

Block, is distinct from any "cognitive, intentional, or functional" notion, rendering it at least conceptually independent from mental processes picked out in these terms (1995 381ff). I will detail below Block's attempts to flesh out and defend this characterization, but first I will present the other kinds of consciousness.

The second kind of consciousness is "access consciousness." A state is access conscious "if it is broadcast for free use in reasoning and for direct 'rational' control of action (including reporting)" (Block 2002b). "Broadcasting" here means that the state is actively available to a number of different psychological systems. Access conscious states influence behavior by flexibly interacting with the goals, beliefs, and desires of a creature. States that are access conscious are "globally accessible" in Baars's sense (1988), or achieve "fame in the brain" in Dennett's terminology (1993). Access consciousness is arguably functionally characterizable. If a state plays the right role in the mental life of a creature, then that state is access conscious.

Block also notes two other kinds of consciousness, "self-consciousness" and "monitoring consciousness." Self-consciousness entails the possession of a self concept and the ability to use this concept in thinking about oneself (1995, 389). Monitoring consciousness (also termed "reflexivity") occurs when a creature becomes aware of one of its own mental states. This occurs when a state is reflected on by a higher-order state of a creature, one that is about another of its mental states. In this way a creature monitors its own mental life (1995, 390). Both these notions plausibly involve cognitive and intentional

processes. We represent ourselves or we represent one of our states when instantiate one of these kinds of consciousness.

Block's claim that p-consciousness is conceptually distinct underwrites his charge that many theorists illicitly avoid the real explanatory problem of consciousness. If he can make good on that claim, what Block terms a "functionalist" theory of p-consciousness cannot work. It will inevitably miss the target due to the nonfunctional characterization of the phenomenon.<sup>7</sup> However, Block argues that this does not obstruct the route to a materialist theory of p-consciousness. Instead, he endorses what he calls a "physicalist" view, on which there is a type-identity between p-conscious states and neural states.<sup>8</sup> P-conscious states just are brain states, despite appearance and intuition to the contrary. Further, there is a decent explanation of our intuitions here: we access our p-conscious states by two distinct conceptual routes. One involves first-person access to the mind; the other involves the theoretical concepts of neuroscience. These concepts refer in very different styles, but there is no bar to claiming they refer to the same thing. This opens the door to the dissolution of Chalmers's hard problem and Levine's gap. However, Block argues that his position is faced with another worry, what he calls "the harder problem of

---

<sup>7</sup> Block includes under this heading the functionalist and representationalist approaches developed by Harman 1990, Dretske 1995, and Tye 1995; the "higher-order" approaches of Armstrong 1968, 1980; Rosenthal 1986, 1997; and Lycan 1987, 1996; the functionalist cum eliminativist approaches of Dennett 1991; Paul Churchland 1981; and Rey 1997; and many others. Any view positing a constitutive connection between p-consciousness and a cognitive, intentional, or functional notion is included.

<sup>8</sup> Block's use of the terms "functionalist" and "physicalist" are not completely standard. Some hold that functionalism view is a form of physicalism (i.e. Lewis, 1994, 291). Other hold that intentionalist views are not functionalist views (i.e. Dretske, 1995). And so on. In this section, I will follow Block's usage.

consciousness" (Block, 2002a). Even if we can avoid the problems of Chalmers and Levine, this issue still lurks. But before presenting Block's physicalism and the harder problem it allegedly generates, I will focus on his characterization of p-consciousness.

Block (with Robert Stalnaker) argues against the use of a priori conceptual analyses in characterizing the data for a scientific theory. He contends that such analyses are never actually produced in sufficient detail and that even the roughest sketches are vulnerable to counterexample. Further, Block contends that a priori analyses are not needed to fix the data. Instead, the data can be fixed by pointing to paradigm examples of the target phenomenon. Uncovering the scientific basis of life, for example, requires focusing on the paradigm living things around here and empirically unpacking the features that explain the common underlying nature of those samples. Furthermore, empirical science has the last word about the data under investigation--if scientific practice dictates a revision of our commonsense intuitions, so be it. Armchair considerations have little role in fixing the data or shaping a scientific theory.<sup>9</sup>

The situation is the same with p-consciousness--we must find ways to "point" to paradigm cases of p-consciousness in order to fix the explanatory data.

---

<sup>9</sup> See Block and Stalnaker 1999. See also Chalmers and Jackson, 2001, for a detailed response. I have addressed some of these issues in the previous chapter, but, to recap, Chalmers and Jackson deny the need for explicit analysis in necessary and sufficient terms. They argue that a priori intensions can be reconstructed by reflection on possible cases, and this is enough to fix the analyses. Block and Stalnaker argue against the existence of a priori accessible intensions, and argue that empirical considerations are always required to fix the data. I attempt to split the difference between the views; I accept the need for the sorts of commonsense analyses that Chalmers and Jackson (following Lewis) require, but I deny that there is any role for the a priori in fixing the data, thus acknowledging the empirical considerations of Block and Stalnaker. See also *Philosophical Studies*, 2004, Davies and Stoljar, eds., on this issue.

However, conscious states are not amenable to standard methods of pointing. One way to get around this problem is to point "via rough synonyms" (1995, 380). We can pick out p-consciousness by employing other terms to pick it out. This, however, is of little value because Block contends that there is no noncircular characterization available. The next form of pointing is to introspect upon the states in question in order to pick out their defining features. Sensations are the prime exemplars of p-conscious states, so we should introspectively reflect on our sensory experience. This perhaps yields the characterization that p-conscious states are states that there is something it is like for the subject to be in. But it does little to support Block's desired distinction--that p-consciousness is not characterizable in cognitive, intentional or functional terms. Arguably, all introspected cases of sensation play a functional role in subject's mental lives. They are always involved in providing information about one's environment or body; further, they are the normal precursor of perceptual belief. In addition, a number of theorists argue that sensory experience is always representational. When we introspect upon our sensations, we always seem to be in a position to note what they represent. Representation can plausibly be cashed out in cognitive, intentional, or functional terms. Thus, introspection does not yield the characterization of p-consciousness that Block is after.<sup>10</sup>

Block then turns to a variety of thought experiments in order to point to p-consciousness. He acknowledges the controversial nature of this method, given

---

<sup>10</sup> See Harman, 1990; Tye, 1995, 2000. See Harman, 1995, for a conceptual (therefore cognitive) view of sensations; Tye 1995, 2000, and Dretske 1995, for intentional views; and Clark 1993, Rosenthal, 1999a, 1999b, for a broadly functionalist view (in terms of quality spaces).

that the proper interpretation of the thought experiments themselves is at issue. However, he believes that, despite the controversy, the cases effectively point to a conceptually distinct phenomenon. Block focuses on Levine's explanatory gap. He writes, "By way of homing in on p-consciousness, it is useful to appeal to what may be a contingent property of it, namely the famous 'explanatory gap'" (1995, 381). He argues that reflecting on possible neuroscientific theories of consciousness leave open the question of why these neurological processes should be accompanied by this p-conscious quality rather than another, or any quality at all. According to Block, we don't have a clue how such theories might answer these questions. Further, this contrasts with our take on current theories of cognition and representation. There seem to be good working paradigms in those cases, in stark contrast with p-consciousness. Block concludes that the gap points to p-consciousness: "...that's the entity to which the mentioned explanatory gap applies" (1995, 382). But this fails to establish Block's distinction. What is it we are thereby pointing to? Perhaps it is a complex functional phenomenon that, as of yet, we have not found an adequate theory. Further, the claim begs the question. Those who deny Block's distinction will also reject the explanatory gap, and for the same reason--they think consciousness can be characterized in cognitive, intentional, or functional terms. To invoke the gap to establish the distinction assumes that this cannot be the case. But that is what is at issue.

Block next offers a number of thought experiments designed to support the distinction directly. He contends that the cases show that p-consciousness is

conceptually distinct from access consciousness, self-consciousness, and monitoring consciousness. Since the other notions are plausibly captured in cognitive, intentional or functional terms, showing p-consciousness's conceptual independence lends support to the idea that the distinction is correct. In particular, he focuses on the independence of p-consciousness from access. Block presents that case of the "super blindsighter" to show the possibility of access without p-consciousness. Blindsight is a condition that sometimes occurs when the visual cortex is damaged. Regions of the visual field will seem to the subject to be completely devoid of visual input. However, when prompted to guess at what is being presented in their "blind field," blindsight subjects are well above chance for certain types of stimuli, like motion, orientation, and simple shape. Block asks us to consider a blindsighted subject who can spontaneously prompt himself to guess about what is present in his blind field and then employ this information in perceptual judgments. Thus, the content of the blind field would be access conscious, because it would play the right role in the subject's mental economy--it would be broadcast for use in speech, for example. Still, we wouldn't think that this "super blindsighter" now enjoyed p-conscious visual experience. Rather, according to Block, it is natural to think that they have access to the information without phenomenal awareness. Access and phenomenality can come apart, it seems.

Block further holds that we can conceive of p-consciousness without access in a creature in which the connections between p-conscious states and the rest of the mind have been ablated. Why think the p-conscious states would

wink out of existence? Isn't it at least possible that such states exist unaccessed? In addition, coming to be aware of a background noise that has been going on for some time potentially indicates the presence of p-consciousness without access. The sound was there, but we didn't access it--it didn't have an impact on a wide range of mental processes. According to Block, prior to the access it is plausible to hold that we were in a p-conscious state, a state involving auditory sensory quality. But until we became aware of it, it went unaccessed. These seem to be cases of p-consciousness without access.

But all of these examples assume the distinction that Block wishes to establish. If we believe that p-consciousness is conceptually linked to cognitive, intentional, or functional processes, then we need not accept Block's interpretation of the thought experiments. In the super blindsighter case, if we believe that p-consciousness is constitutively tied to access, then, if the right access relations are really present, it follows that the super blindsighter has p-conscious states. Further, the details of the access matter. Many views reject the claim that the ability to self-cue guesses is sufficient for access. When the super blindsight case is restated with that in mind, Block's conclusion becomes less plausible. There may be some access present, but not the sort of access sufficient for phenomenal experience.<sup>11</sup> Finally, it is not clear that the thought experiments indicate anything about our "pretheoretic" notion of consciousness. By the time we come to grasp the subtleties of these imagined cases, we lose our grip on the initial data we wanted to pin down. How do we know what an

---

<sup>11</sup> For a related discussion, see Siewert 1998, and the commentaries on his work on the PSYCHE website, <http://psyche.cs.monash.edu.au/symposia/siewert/index.html> (Siewert 2004), especially the commentaries by Lycan and Carruthers.

imaginary brain-damaged subject will experience? How are we to tell if ablating an animal's brain leads to this or that experiential result? Our folk concept does not extend out to these cases. It is only if we are already sure that they are distinct in the way that Block claims that we will find his readings appealing. But that goes beyond mere "pointing" and into full-blown theorizing. Thus, if it is to provide a datum that a materialist theory of mind must explain, Block's characterization of p-consciousness requires further support.

So, where can we turn for that support? In his 1995 paper, Block claims that "[t]hough I believe that functionalism about p-consciousness is false, I will be trying not to rely on that view" (1995, 381). But this is the root of the issue. So it is useful to look back on Block's previous rejection of functionalism. In arguing for the failure of functionalism, it may be that Block provides support for the claim that there is a kind of consciousness that defies characterization in cognitive, intentional, or functional terms.

In his 1978 paper "Troubles with Functionalism," Block presents several cases designed to undermine a functionalist theory of "qualia," the qualitative aspects of conscious experience. Qualia are paradigm cases of p-conscious properties. The thought experiments are very much in the mold of those already presented. One involves imagining the nation of China "wired up" by radios in order to instantiate a functional analogue of human psychology. All the inputs, outputs, and intermediate functional states could be mimicked (in principle, it is claimed) by this organized mass. But, Block contends, it is absurd to think the mass itself would thereby be in conscious qualitative states. But functionalism is

committed to this position, rendering it absurd as well. A second related case focuses on a broader form of functionalism, one only committed to the equation of mind with the actual functional organization of our brains, rather a more abstract functional-role characterization. This is the case of the "homunculi head," where our neural organization is reproduced by tiny sentient creatures who manipulate the requisite connections and processes. But again, Block holds that it is counterintuitive to believe that such a functional system would possess qualia. The presence of the sentient beings in the middle of the process helps bring out the problems functionalism has in explaining our conscious qualitative states.

However, such cases do no more than restate the intuitions that drive the claims in the 1995 paper. Further, it is again open to the functionalist to argue that such creatures would be p-conscious, despite our intuitions. Such intuitions may simply represent our pretheoretic ignorance about the nature of our conscious minds, rather than the conceptual distinction Block is after. Indeed, neurons are characterized fully in functional terms in neuroscience. However, simple reflection on our neurons fails to reveal how they could instantiate p-consciousness, either. But Block accepts that they do. So why not think that the situation might be the same in the problem cases? Block argues that we know that we are conscious, so we have reason to believe that neurons underwrite p-consciousness in our case. But we have no parallel reason for such a belief with non-brain functional systems. But that misses the point. We may not recognize that what matters for p-consciousness is the functional arrangement of our

neurons, rather than their constitution. Why should we be able to tell this? And if that is the case, than a system that recapitulates that organization will also possess p-consciousness, despite the intuitions.<sup>12</sup>

For all of my criticisms, however, it may still seem Block is quite correct that p-consciousness is distinct from any cognitive, intentional, or functional notion. There must be something underwriting the pull of the thought experiments. Isn't that alone enough to motivate the conceptual distinction? It may seem to follow from the definition alone that p-conscious states are distinct. Block characterizes them as states that there is something it is like for the subject to be in, and this maps directly to Chalmers's problematic notion. Perhaps this is enough to establish the independence that Block seeks. If a state is characterized in terms of "what it's like," what does that have to do with cognition, intentionality, or function?

However, as noted in chapter 1, there is more than one way to unpack the "what it's like" locution. One focuses on the front end of the clause, on the something that it's like for the subject. This suggests a property glowing with consciousness, one that by itself carries the problematic feature that cries out for explanation. Perhaps incidentally it is for a subject, but the quality itself floats free from cognition, intentionality or function. But the other way of unpacking the phrase puts the emphasis on the "for the subject" element. In this case, if the state isn't registered by the subject at all, it is wrong to call it a conscious state.

---

<sup>12</sup> An additional support of support for Block's distinction might be his "inverted Earth" claim from his 1990 paper of that name. But this issue is largely the same. It is open to the representationalist to either embrace Block's conclusion, or to argue for a more restricted version of the thesis. See Harman 1990, 1999; Dretske 1995; Tye 1995, 2000.

This second reading of the phrase is stressed by David Rosenthal (2002b). He argues that sensory states (the "something") that are in no way for a subject fail to qualify as conscious states. If a subject is in no way aware of the state she is in, it is not intuitively a conscious state. The presence of the subjective element is what matters for consciousness. Such a reading is in line with our commonsense characterization of consciousness. States that we are in no way aware of are not intuitively conscious states. This seems as good a way to characterize states that there is something it is like for the subject to be in as Block's.

Block acknowledges that there are alternate ways to fix the data, but he argues that his distinction presents a theoretically useful way to interpret empirical results. This, he feels, provides a modicum of empirical support for his position. If it provides a useful way of clarifying and interpreting psychological experiments, the distinction gains legitimacy. Indeed, Block argues that even if his distinction isn't unambiguously supported by empirical results, "we can get many convergent though flawed sources of evidence--[and] so long as the flaws are of different sorts and uncorrelated--we will have a convincing case" (2001, 217, emphasis in original). He offers several examples to make his point.

One is the condition Block labels of "aerodontalgia." Pilots in World War II flying in unpressurized planes occasionally complained of feeling pain in places where they had undergone dental work, despite the fact that this work was done under anesthetic. It turned out that only that work done under general anesthesia produced the recalled pain; work done under local anesthesia failed

to produce the experience. Block hypothesizes that this indicates the presence of phenomenal consciousness (pain) without either access or reflexive awareness. The pain must have been present in order to lay down the pathways activated later in flight. General anesthesia therefore must only block access or reflexivity, rather than blocking the p-conscious pain state altogether. This seems to be a useful way of describing the case, and one that points to distinct phenomena to be studied. The distinction between p-consciousness and cognitive, intentional, and functional notions seems fruitful.

However, rival interpretations can handle the case with equal ease. The advocates of access or reflexivity can argue that in the general anesthesia case, the pains present at the time of the dental work either failed to be accessed or were not the objects of reflexive awareness. And, therefore, the pains were nonconscious--there was nothing it was like for the subject to experience the pains. If access or reflexivity are constitutive of conscious states, then nonconscious sensory states are an accepted possibility. Furthermore, there is good evidence that folk psychology licenses nonconscious sensory states; additionally, they have wide acceptance in empirical work. People regularly speak of a headache that went on all day, though at times the pain wasn't conscious. Also, sensory states that occur very quickly are often termed subliminal perception. It is accepted usage to call them unconscious sensory states, both in commonsense talk and in scientific research.

Block is only in a position to rule out nonconscious states if he has evidence from some additional source that such states are impossible--

otherwise, he is question begging in the current context. If that source is our commonsense, pretheoretic notion of consciousness, then at best he gets a standoff. Some cases pull the usage one way, some the other, and neither gains the upper hand. Block runs through a number of other empirical cases, from commonsense reflection on the experience of hearing a jackhammer, to priming and perception experiments in psychology. However, the line of response is the same as in the aerodontalgia case. The other side has a reasonable interpretation in terms of nonconscious or partially accessed or reflected upon states. The cases are a standoff. Thus, both sides gain the same "convergent though flawed sources of evidence." Block's view gains no advantage from empirical data.

At this point we are left only with more general theoretical considerations to separate the views. These include the usual theoretical "virtues" of predictive power, simplicity, scope, familiarity of principle, fruitfulness, etc. (see Quine 1976; Quine and Ullian, 1970). But in that case, it is arguably the access or reflexivity theories that gain the upper hand. These theories integrate p-consciousness with the rest of scientific psychology and neuroscience and they avoid both the explanatory gap and the hard problem. The main challenge for such views is saving, as much as possible, our pretheoretic intuition concerning consciousness. But this seems a more desirable theoretical situation than being saddled with the hard problem or the gap. And there is an additional consideration that tells against Block's distinction in this context: Block's own theory of consciousness, which embraces the independence of p-consciousness,

leads to what he terms the "harder problem of consciousness." Thus, even if the distinction is granted, substantial explanatory problems are incurred. In the following section, I will present Block's harder problem; however, I will argue that it is not significantly different from Levine's explanatory gap. I conclude that Block's theory fails to explain consciousness, and therefore, his distinction is not theoretically fruitful. Thus, we have no compelling reason to accept that p-consciousness is conceptually distinct from any cognitive, intentional, or functional notion.

### **2.2.2 Block's Identity Thesis and the Harder Problem of Consciousness**

To review, Block argues against any a priori analysis of p-consciousness in cognitive, intentional, or functional (or any other informative) terms. Further, he contends that p-consciousness is conceptually distinct from any such notion. Still, he holds that physicalism is true, and therefore, that p-consciousness is ultimately a physical phenomenon. He argues that p-conscious states are identical to neural states. The identity is justified on grounds of ontological parsimony and explanatory power. We are thus entitled to posit the identity, even in the absence of any deductive link between the two phenomena, contrary to Chalmers's arguments.<sup>13</sup> Furthermore, identity claims do not allow for meaningful explanations. Thus, the lack of an informative explanation of the

---

<sup>13</sup> Levine, like Chalmers, holds that successful explanations must deductively entail their explanans. However, he does not hold that identities are justified deductively, as Lewis, Chalmers, and Jackson argue (as did Levine in his 1993). Levine, like Block, holds the identities are posited because of their explanatory and ontological value. Once the identity is in place, we should be able to construct the requisite deduction needed for explanation. See Levine 2001, Chapter 3, section 2.

consciousness/neuronal identity is not an objection to the position. It simply follows from the fact that identities are not informative in general.

So, how do we arrive at a scientific understanding of p-consciousness, according to Block? First, the data for the theory is fixed by picking out paradigm cases of p-consciousness--a sensory experience of red, for instance. Then, using brain imaging and other empirical methods, we isolate the "neural correlates" of p-consciousness, the neural states that reliably co-occur with p-conscious states picked out introspectively. We then seek to corroborate such evidence with dissociative disorders and other clinical data. When enough evidence is in hand, we posit an identity claim between the experiential state and the neural correlate, in the interest of explanatory power and ontological parsimony. So, for example, it might be that visual sensations of red are correlated with a certain type of cell firing in areas V1 and V5 of the visual cortex. According to Block, we are justified in claiming that visual sensations of red just are cells of that type firing in the relevant areas. We thereby close off questions of why these two phenomena are always correlated, and reduce the number of distinct entities in our ontology. Finally, all this is achieved without recourse to a priori analysis, contra Chalmers and Jackson.

Block also argues against the legitimacy of apparently open questions used by Levine to support the gap. Identities do not have explanations, according to Block. Asking for an explanation of an identity is just to ask how something could be what it is. Still, we can ask for reasons to believe that we've found a true identity, rather than a mere correlation of distinct entities. And Block

admits that there is considerable mystery about how things that seem as diverse as sensations and neural states could be identical. The solution, according to Block (in agreement with Loar, Papineau, Perry, and others, and even Levine in a qualified form), is to recognize that we can pick out the same thing under more than one concept. Block argues that our first-person access to p-consciousness is mediated by "phenomenal concepts," concepts that achieve their reference by somehow instantiating the very phenomenal states they are about. Because of the involvement of p-consciousness in phenomenal concepts, they seem to pick out something radically different from the referents of theoretical concepts employed in neuroscience. Those concepts are in the descriptive language of scientific theory. Still, this does not preclude the possibility that both concepts share the same referent. Thus, according to Block, we "replace a dualism of properties with a dualism of concepts" (Block, 2002b). This is acceptable from a physicalist perspective, so long as the phenomenal concepts themselves do not require nonphysical elements to do their work.

The dualism of concepts helps dispel the appearance of an explanatory gap, according to Block. The distinctive nature of our first-person access to consciousness explains why we are left with lingering questions. But given the truth of our identity, such questions are illegitimate. While Block acknowledges that the details of the identity are still sketchy (particularly on the neural side), he contends that we have a clear path towards closing the gap. Thus, both the metaphysical hard problem of Chalmers, and the epistemological explanatory gap of Levine fall to Block's version of physicalism.

However, there is another problem lurking for Block's brand of physicalism. The "harder problem" of consciousness arises from the tension between several theoretical commitments of the view (Block, 2002b). First is a commitment to what John Perry (2001) calls "antecedent physicalism," the idea that physicalism is the default reasonable position to take concerning the metaphysics of mind (here "physicalism" just means "materialism" in Levine's sense). Second is a belief in "naturalism," a commitment to the methods of science and to the application of empirical methods in theorizing about consciousness. Third is a dedication to "phenomenal realism," which consists in a rejection of an a priori analysis of p-consciousness into cognitive, intentional or functional terms--this falls out of his characterization of p-consciousness. Phenomenal realists thus hold that consciousness can't be unpacked into the more tractable functionalist notions of mind.<sup>14</sup> Finally, Block's view rejects skepticism about "other minds." It is taken to be beyond doubt that other humans have minds with p-conscious states, while rocks do not, though there is no functionalist justification of this claim. Instead, it is rooted in beliefs about both our functional makeup and our flesh and blood constitution. Other minds skepticism is not, according to Block, the basis of the harder problem.

Given these commitments, Block argues that an epistemological worry arises concerning the possibility of alien or robot consciousness. He considers the case of a robot, Star Trek's Data. Data's behavior is similar enough to our

---

<sup>14</sup> Block calls the views that accept such analyses "deflationists" about p-consciousness; such views hold that p-consciousness is ultimately nothing more than one of the better understood facets of the mind in different attire. See Block, 2002a.

own to legitimately raise the question of whether or not he has p-conscious states. However, Data is made of stuff utterly unlike our own flesh and blood. This removes any obvious parallel to our neural correlates of p-consciousness. In our own case, we are able to correlate p-consciousness with neural states by employing our first-person access to the phenomenon. But we are not sure if Data is p-conscious at all. Why think the causes of his verbal "reports" are p-conscious states, rather than nonconscious analogues? In our case, first-person access is a crucial step in identifying the neural correlates of p-consciousness. But we cannot access Data's states in this way. Further, any reliable functional mark of p-consciousness is blocked by the endorsement of phenomenal realism. This removes any third-personal way of settling the issue. Thus, when faced with Data, we are at a loss. Is he made of the right stuff to support p-consciousness? And how could we possibly figure this out without being Data?

We are left in the following situation, according to Block. We have reason to think that physicalism and phenomenal realism are true. But our naturalism dictates that we must view the possibility of Data's p-consciousness as an open, empirical question. Yet such a question seems unanswerable in principle. Block admits that this is not a full-blown paradox for the physicalist, but he insists that it presents an uncomfortable epistemic tension in the view. A question that ought to be empirical and open is arguably beyond our means to settle. Yet it seems a fair and reasonable question, one that a theory of consciousness ought to address. This is the "harder problem of consciousness."<sup>15</sup>

---

<sup>15</sup> Block adds considerable detail to the argument by stressing the epistemic strength of various belief commitments that accompany the view. He distinguishes between reasons for believing,

### 2.2.3 Block and Levine

Block and Levine occupy very similar regions of theoretical space on the problem of consciousness. Both hold that materialism is the default metaphysical position on the mind, given its ability to account for the causal efficacy of the mental and to deliver a unified scientific picture. Further, both reject the use of a priori conceptual analysis and argue that conceptual content, including content about consciousness, is fixed using a posteriori methods. Both also agree that phenomenally conscious states are conceptually distinct from functional processes, and this presents an in-principle obstacle to a broadly-construed functionalist theory of consciousness. This creates an apparent explanatory gap between the physical and the phenomenal; indeed, the gap seemingly defines the phenomenon. Finally, both agree that the epistemic gap carries no ontological weight, because the gap can be explained in terms of distinct concepts picking out the same referent. This clears the way for a physicalist identification of p-consciousness and neural states.

However, the two differ over the implications of the explanatory gap for a satisfying theory of consciousness. According to Block, the two-concepts line offers a route to closing the gap. When viewed in proper conjunction with the identity of p-conscious states and brain states, the presence of distinct concepts explains why we seem to have meaningful questions. Once we grasp the

---

grounds for rational belief, and certainty, and employs these notions throughout. I have simplified the picture in the interest of clarity, and because it will not affect the main thrust of Block's argument nor my comments on it. For a detailed reconstruction of Block's argument, see McLaughlin 2003.

divergent ways we access the mind, our apparently meaningful concerns are exposed as empty. We may still feel the explanatory itch, but that is simply a reflection of how deeply we sense the difference in concepts. But it does not signify meaningful open questions indicative of a real explanatory gap.<sup>16</sup>

Levine counters that there are meaningful open questions in the consciousness case, due to the substantive and determinate nature of our access to consciousness. Levine, recall, argues that we access our conscious states from the first-person perspective by employing phenomenal concepts with substantive and determinate thick MOPs. Because we have such apparently direct access to what we are referring to, we fail to see how p-conscious states could just be brain states. It is not simply a matter of ignorance; rather, we know p-consciousness in a way that holds open valid questions. This differs greatly from other identity claims, so Levine concludes that identities involving substantive and determinate MOPs on one side and thin MOPs on the other leave meaningful open questions. This signifies a gappy identity, and hence an explanatory gap.

Block in turn rejects Levine's explanation of gappiness. He argues that there are identities involving substantial and determinate MOPs that nonetheless fail to display gappiness. He writes

Consider that the mode of presentation of a sensation of a color can be the same as that of the color itself. Consider the identity 'Orange = yellowish red'. Both modes of presentation involved in this identity can be

---

<sup>16</sup> Cf. Papineau 1995; 2002 chapter 4.

as substantive as those in the putatively "gappy" [mind-body] identity..., yet this one is not "gappy" even if some others are. To get an identity in which only one side is substantive, and is so a better analogy to the mind-body case, consider an assertion of 'orange=yellowish red' in which the left hand concept is phenomenal but the right hand concept is discursive (Block, 2002b, xx).

Block thus concludes that Levine's explanation of gappiness fails. There are cases where substantive and determinate MOPs do not result in gappy identities; therefore, a better explanation of the lingering questions in the consciousness case is the extreme distinctiveness of our concepts, rather than a revealing sort of first-person access. The residual questions remain illicit.

But Block's claim is not convincing. The fact that MOPs of color sensations and colors can be the same does not provide the needed contrast. If orange=neural state x, then we have Levine's gappy identity. But if we consider orange=light of 500 MHz, where orange is presented with a phenomenal MOP, we get a similar gappy identity. It is one of the central difficulties of color theory to explain how it could be that a wavelength just is a color, given the way colors appear to us. Wavelengths just don't seem to be the right sort of thing to possess the qualitative aspects that define color. In fact this problem has driven some theorists to be dualists, and others to be eliminativists, about color. And it is plausible because we possess a substantial and determinate MOP of orange, because we have seemingly direct access to the quality itself, that we find these identities mysterious.

Block's claim concerning the identity of orange and reddish-yellow, where reddish-yellow is taken "discursively," is also unconvincing. What exactly is being identified here? Let us allow that one can possess the purely discursive concept "reddish-yellow" in the absence of any phenomenally grounded knowledge of the color (otherwise, the identity would be non-gappy due to the presence of substantial and determinate MOPs on both sides of the identity). This would plausibly involve knowing about reddish-yellow in relational terms, for example in terms of its location in a relational quality space. Thus, orange=point  $x$  in the quality space, the space in between red and yellow picked out in the same fashion. But this is also plausibly a gappy identity. How could the sensation of orange just be a location in an abstract quality space? Such a claim unpacks the sensation in relational terms, or in the cognitive, intentional, or functional terms that Block rejects. It is not open to a "phenomenal realist" to embrace this view. And furthermore, it leaves open exactly the questions that Levine claims are indicative of a gap (see Levine 2001, sections 4.2-4.3). Block has failed to provide a counterexample to Levine's position. Thus, Levine's explanation of gappiness holds, and Block must find a way either to address the remaining meaningful questions or to explain them away. If he cannot, the explanatory gap remains as a worry.

I argued in the section on Levine's gap that what really demands explanation is the appearance of the gap. To put things into Levine's terminology, what demands explanation is our apparently substantial and determinate access to seemingly intrinsic qualities of experience. If we can

provide such an explanation, it is open to us to reject the gap as illusory. In a sense, this is what Block wishes to do with his two-concepts line. Levine also has sympathy with the two-concepts line, but he argues that it fails to close the gap because we do not understand how the presence of the phenomenal state in the phenomenal concept itself can make a cognitive, referential difference. He writes

It does seem as if the very property of reddishness is somehow present in the concept, making a cognitive contribution that endows the content with genuine substance and explains the gappiness of the identity. But how do we explain that on a physicalist model? How does a property referred to by a mental representation get cognitively incorporated into the representation in the way it seems to with phenomenal concepts and properties? ... What emerges...is that the explanatory gap is intimately connected to the special nature of phenomenal concepts. ...But then we have the problem of providing an explanation in physicalist terms of that very specialness, and we don't seem to have one (2001, 86).

Simply saying that a p-conscious state is in a phenomenal concept is not sufficient to explain how this special mode of access works. And in the absence of a more detailed model, it is fair to say that legitimate open questions remain.

Therefore, the gap is still a worry for Block.<sup>17</sup>

---

<sup>17</sup> I will have much more to say about the "two concepts" model in chapter three, where I present and criticize the approach in detail. Block's comments in his 2002b are, by his own admission, sketchy. He refers favorably to the work of others in that context, particularly the work of Brian Loar. I address Loar's proposal in chapter 3 below.

And that is not a surprise. Levine insists that the issue at the root of Block's harder problem, the puzzle of robot or alien consciousness, is just a facet of the explanatory gap, and not an independent worry. According to Levine,

Another way to see a manifestation of the explanatory gap is in our deep puzzlement over the question of attributing conscious experiences to creatures somewhat different from ourselves... What we lack is a principled basis for determining how to project the attribution of conscious experience. I submit that we lack a principled basis precisely because we do not have an explanation for the presence of conscious experience even in ourselves. We know, perhaps, or have good reason to believe, that its presence is due to something about our physical constitution. But without an explanation of how our physical constitution gives rise to consciousness, we can't use that knowledge as a basis for determining what else has it... The fact that we don't know what to look for is... a manifestation of explanatory ignorance (2001, 77-78).

If he is correct, than arguably there is no substantial difference between the two positions. There is just the explanatory gap, though it manifest in a number of guises.

We can see the collapse of the harder problem into the gap by noting that a solution to one entails a solution to the other, and vice versa. First, if we had a solution to the explanatory gap, we would be in a position to extend our concept of consciousness and dissolve the harder problem. A solution to the gap would answer how the physical brain achieves phenomenal consciousness, what it is

about particular neural tissue that makes it a phenomenally conscious state. That kind of understanding would provide a principled basis for the relevant extension. If the range of extension were still a mystery, we would still be in a position to ask what it is about our neural tissue that makes the crucial difference, and therefore, the gap would still be open.

In the other direction, any solution to the harder problem would entail a solution to the explanatory gap in our own case. Again, we would then be in a position to determine in a principled way if an arbitrary alien with a radically different constitution possessed phenomenally conscious states. To do that, we would need to know what sort of relevant connection there is between constituting states and phenomenal states. But that would offer the same sort of guidance in our own case, closing the explanatory gap. Thus, the two problems do not come apart. Levine is correct that the problem of robot and alien consciousness is just a manifestation of the explanatory gap, and his position and Block's position do not diverge.

To conclude, Block's view is not significantly distinct, in the final analysis, from Levine's position. Both are physicalists whose positions fail, largely for the same reasons, to close the explanatory gap. Block's view does not present a new, "harder" problem of consciousness; rather, it accentuates a previously identified manifestation of the explanatory gap. The main challenge posed by the gap, as argued in section 2.1.2, is to account for the apparently substantial and determinate access we have to our conscious states. Furthermore, Block's characterization of p-consciousness is not the best way to pick out the

phenomenon, and upon scrutiny, his claims lack support. The way remains open for a characterization of p-consciousness in cognitive, intentional, or functional terms, and thus to a reductive explanation of consciousness. In the concluding section of this chapter, I will lay out the residual explanatory task remaining after our criticisms of Chalmers, Levine, and Block. I will argue that explaining our access is central to opening the route to a satisfying materialist theory of consciousness. However, that explanation of access must also account for the appearance of seemingly intrinsic features of consciousness.

### **2.3 Conclusion: What's the Problem?**

After considering the efforts of Chalmers, Levine, and Block to spell out a special explanatory problem of consciousness, we are left with the following results. At bottom, all three theorists hold that consciousness does not appear to be functional, from the first-person point of view. Worse, it seems to them that there isn't an illuminating connection between consciousness and any relational notion, any notion specifiable in terms of an entity's relations to other things. This appears to cut off consciousness from the sorts of connections ordinarily used by science to pin down and explain a target phenomenon. Further, following Levine, we seem to have direct access to our conscious states, and what's more, we seem to be in contact with intrinsic features of consciousness. We seem to be in direct contact with nonfunctional aspects of the conscious mind.

This is how it appears to us, according to these three theorists. However, all three attempted to move beyond these appearances to make substantial

claims about the nature of consciousness. Chalmers argued that the appearances underwrite our phenomenal concept of consciousness, and given the nature of our concepts, this leads irreducibility and the failure of supervenience. Levine argued that the appearances undermine the possibility of a materialist explanation of consciousness--we can tell from the appearances that the usual type of materialist theory is doomed to failure. And Block argued that the appearances support the presence of a distinct concept of consciousness not characterizable in cognitive, intentional or functional terms. This blocked the possibility of a functionalist theory of consciousness, forcing us to posit an identity claim which ultimately failed to shed light on how it is that the states of a physical brain could be conscious states.

I argued that all three of these claims fall short. Given our commonsense approach to fixing the data, we require successful argument to move from the appearances to the underlying nature of consciousness. Folk psychology alone does not license this move; independent justification is needed. However, under scrutiny, that justification was lacking. Chalmers's model of reductive explanation failed to reliably rule out the possibility of conceptual change in the face of theoretical advance. Levine and Block also failed to provide the needed support for their characterization of a nonfunctional kind of consciousness. At best, the three theorists established that it appears to us that consciousness is nonfunctional.

This opens the door to a materialist explanation of consciousness. If we can explain why it appears to us that consciousness is nonfunctional in a way

that does not entail that consciousness really is nonfunctional, we can then employ the usual methods of theory-building in psychology and neuroscience. If there only appears to be intrinsic features of consciousness, then a relational theory is fully adequate. The apparently intrinsic features can be explained away as a product of our access mechanisms. A satisfying explanation of those mechanisms explains why things appear as they do to common sense, and why we are at times swayed by the intuitions cited by Chalmers, Levine, and Block. But it would not require us to take those appearances as accurate. If such an explanation can be developed, there is no barrier to a fully satisfying materialist explanation of consciousness.

We must therefore pin down the conditions of adequacy on a model of our first-person access. Such a model must explain both the seemingly direct nature of our access, and the seemingly intrinsic features we seem to access directly. In the next chapter, I will provide more detail about what is required of such a model. Then I will present in detail and criticize the "two concepts" approach mentioned in conjunction with Block. This approach has become the standard move in explaining the intuitions that create resistance to materialism. However, I will contend that under scrutiny, the view falls short. In chapter 4, I will present a rival model, one that successfully explains the appearances while remaining firmly within the materialist framework.