

Phenomenal Consciousness and the Explanatory Gap

Karol Polcyn

There seems to be a mystery as to how phenomenal consciousness could arise out of physical processes or states. One intuition that gives rise to this mystery is that it is conceivable that all the physical processes are as they are and that consciousness is missing¹. If consciousness were dependent on physical processes, this would not be so conceivable. For if consciousness were dependent on physical processes, those processes would explain why conscious states feel the way they do and hence the relation between phenomenal states and physical ones would be necessary. Another intuition that gives rise to the mystery is that it is hard to identify the property that would explain consciousness. There are some powerful reasons to think that the identification of such a property is impossible for us due to our cognitive limitations².

All those intuitions support one version of the problem of the explanatory gap. Consciousness simply resists explanation in physical terms, that is, we cannot explain one crucial aspect of conscious states, why it is like what it is like to be in those states, and therefore it is hard to see how conscious states could be *generated* by physical processes. One way to remove this mystery would be to deny that consciousness is generated by physical processes or that, in other words, physical processes give rise to consciousness. We might think instead that conscious states and physical states are *literally identical*. Indeed, we seem to have all the empirical evidence for the truth of the identity: the behavioral effects that we associate with conscious states are the effects of certain physical states. In what follows, however,

¹ This argument is pressed by Chalmers [1996]. Of course, whether or not it is conceivable that there should be a world physically identical to our world and deprived of consciousness is a matter of controversy. Physicalists deny the conceivability of such a world. For an extended discussion of this point see Chalmers [2002]. It is also worth pointing out that there are different senses in which we might say that something is conceivable and that corresponding to those different senses there will be different relations between what is conceivable and what is possible. For more on this, see Gendler and Hawthorne [2002].

² See McGinn [1989].

I will argue that we do not understand how phenomenal and physical states could be literally identical. Despite the fact that we have all the empirical evidence for the truth of the identity, the identity does not seem to be fully intelligible.

My argument to this effect will build again on the impossibility to explain consciousness in physical terms. As it will become clear, this difficulty bears not only on the question whether consciousness depends on the physical but also on the question whether consciousness is identical with the physical. That is not obvious and is, in fact, a matter of strong controversy in the current literature. Many philosophers simply deny that our inability to explain consciousness in physical terms undermines the coherence of assuming that psychophysical identity is true. I agree with those philosophers to some extent, that is, I agree that despite the fact that we lack the explanation of consciousness, we still have all the empirical evidence for the truth of psychophysical identity we need. However, I will argue that even though the explanatory gap does not raise the issue of the justification of psychophysical identity, it raises the issue of the intelligibility of that identity. My point is this. Assuming that psychophysical identity is true, physical states can be *described* phenomenally, as states that feel a certain way, and assuming that their phenomenal character (its being like this or that to be in any given state) cannot be explained in physical terms, there is simply no answer to the question as to why they should be described phenomenally in any particular way or why they should be described phenomenally at all³. It is in this sense, I claim, that psychophysical identity remains unintelligible.

1. The explanatory gap

Here is the problem of the explanatory gap as it was introduced by Levine [1983, 1993] in the context of the question whether consciousness is identical with

³ It is often assumed that the concepts of conscious experiences are not descriptive concepts since they do not pick out their referents via the description of some properties different from those referents. Nevertheless, since those concepts conceive of their referents as being a certain way, it is natural to assume that there are certain modes of presentation (phenomenal modes of presentation) associated with those concepts. Here I am using the term “phenomenal mode of presentation” as synonymous with “phenomenal description”. I am assuming, of course, that phenomenal descriptions, descriptions of the form “this is what it is like”, are very unique descriptions: they cannot be understood unless one has actually had the experiences these descriptions are true of.

physical properties. The problem stems from the recognition of some crucial differences between psychophysical identity, on the one hand, and theoretical identities discovered by science, on the other. Some key differences between those identities were already observed by Kripke. According to Kripke, whereas both sorts of identities appear contingent, the apparent contingency of psychophysical identity cannot be explained away as a semantical illusion (that is, as a misdescription of what is really conceivable). The problem of the explanatory gap, on the other hand, has to do with the fact that psychophysical identity does not have the explanatory force of other theoretical identities. As it will become clear, though, and as Levine himself points out, this explanatory deficiency of psychophysical identity has its roots in the fact that the apparent contingency of psychophysical identity cannot be explained away.

What is the difference between the explanatory force of psychophysical identity and that of other theoretical identities? Theoretical identities discovered by science are explanatory in the sense that they provide the complete explanation of the manifest properties of the relevant natural kinds. Psychophysical identity, on the other hand, is *not fully* explanatory for it does not explain all the manifest properties of conscious experiences. While it affords an explanation of the causal role of experiences, it leaves their phenomenal character unexplained.

To see the difference, let's contrast the alleged identity of pain and c-fiber stimulation with the identity of water and H₂O. The latter identity explains all the manifest properties of water, such as its being liquid at room temperature, its freezing and boiling points, etc. (in short, watery properties). However, the identity of pain and c-fiber stimulation is not fully explanatory for it does not explain why it is like what it is like to feel pain. The identity may explain why pain has its characteristic behavioral effects but the phenomenal character of pain is left unexplained.

As Levine points out, the basis of this explanatory gap is conceptual. That is, the explanatory gap arises because for any given pair of physical and phenomenal states it is conceivable that while the physical state occurs, the phenomenal state does not. So in the case of the identity of pain and c-fiber

stimulation, the fact that this identity leaves the explanatory gap has its root in its being conceivable that while c-fiber stimulation occurs, there is nothing it is like to be in pain. Now, the reason why such conceivability intuitions lead to the explanatory gap has to do with the fact that explanatory relations are conceptual in the sense that the phenomenon to be explained is entailed a priori by the explaining phenomenon. Consider again the identity "Water = H₂O" and suppose we want to explain why water boils at 212°F. The explanation will appeal to some causal mechanism that involves the specific behavior of H₂O molecules at 212°F. There is no room to specify the details of that mechanism here. The key point is that assuming that the explanatory mechanism is in place (and assuming that the physical and chemical laws are as they are) it will follow a priori that water should boil at 212°F. Thus, if we assume that the mechanism is in place, it will be inconceivable that water should not boil at 212°F. Similarly, if we could explain why pain feels like pain in terms of some causal mechanism involving c-fiber stimulation, it would be inconceivable that there should be c-fiber stimulation without there being anything it is like to feel pain. So since it is conceivable that c-fibers stimulation might occur without there being the feeling of pain, we are forced to assume that the identity of pain and c-fiber stimulation does not explain the phenomenal character of pain and that, in this sense, it leaves the explanatory gap.

What is the significance of the explanatory gap? Levine himself thinks of the explanatory gap as an epistemological rather than a metaphysical problem. According to him, the existence of this gap does not show that psychophysical identity is not true. As we just saw, the explanatory gap results from its being conceivable that psychophysical identity be false. We might think that this creates a *prima facie* difficulty for assuming that psychophysical identity is true since the identity, if true, is true necessarily. But Levine assumes that conceivability does not imply possibility and hence that despite its being conceivable that physical and phenomenal states should come apart, the relation between them could still be necessary. We may not be able to explain why they are necessarily related but then, again, this would only be the reflection of our epistemic situation, that is, it

would only be the reflection of the fact that we find the falsity of psychophysical identity conceivable.

Now, whether or not Levine is right about the lack of implication from conceivability to possibility, the existence of the explanatory gap does seem to create a difficulty for physicalism. For the gap implies that psychophysical identity is *not fully intelligible*. Levine himself seems to encourage this line of thinking. Here is the relevant quote:

There seems to be nothing about C-fiber firing which makes it naturally “fit” the phenomenal properties of pain, any more than it would for some other set of phenomenal properties. Unlike its functional role, the identification of the qualitative side of pain with C-fiber firing (or some property of C-fiber firing) leaves the connection between it and what we identify it with completely mysterious. One might say, it makes the way pain feels into merely a brute fact⁴.

The question of intelligibility that is at issue here should be distinguished from the question of justification. For the question of the intelligibility of psychophysical identity arises even though we have all the empirical evidence for the truth of that identity.

Why should that be so? The reason is that whereas the justification of psychophysical identity can be based solely on our ability to explain in physical terms the causal role of experiences, this sort of explanation falls short of making the identity fully intelligible. As for the justification, the point seems obvious. Conscious states seem to be identical with those physical states that cause the behavioral effects that we attribute to conscious states. So, for example, assuming that c-fibers firing causes the behavioral effects that we attribute to pain, we should be justified to conclude on this basis that pain and c-fibers firing are identical.

Levine himself acknowledges the force of this intuition. As he says: “It is precisely on the grounds that a particular physico-functional property can explain the «behavior» of qualitative states that we would endorse an identification

⁴ Levine [1983] p. 357.

between a particular quale and that property"⁵. According to Levine, however, something more is required to make psychophysical identity intelligible. Apart from explaining the causal role of experiences, we need to make intelligible their phenomenal character. That is, we need to make intelligible why it is like what it is like to undergo any given experience. For example, in order to make intelligible the identity of pain and c-fibers firing, we need to make intelligible why it is like what it is like to feel pain. That task, obviously, cannot be accomplished simply by explaining the causal role of pain. It is one thing to explain the causal role of pain and another to explain its phenomenal character.

The corresponding challenge does not arise in the case of other theoretical identities, such as the identity of water and H₂O. The concept of water is a causal role concept and hence once we have seen that that causal role is played by H₂O, there isn't any further question as to why water and H₂O should be identical. Once we have seen that the causal role of water is played by H₂O, the identity of water and H₂O follows a priori, simply by our understanding of the concept of water. To ask in those circumstances the further question of why water should be identical with H₂O would be a sign of a conceptual confusion.

The reason why things are different in the case of psychophysical identity is that phenomenal concepts are not causal role concepts.⁶ For example, the concept of pain does not conceive of pain as playing such and such causal roles. Instead, it conceives of pain as having a certain phenomenal character that is conceptually distinct from any causal roles. That conceptual distinctness simply follows from the fact that one can conceive of a given state as playing a given causal role associated with pain and yet not constituting any qualitative experience of pain or constituting an experience that is qualitatively distinct from the experience of pain⁷. Now, the fact that the concept of pain is not a causal role concept guarantees that even after we have explained the causal role of pain in terms of c-fibers firing,

⁵ Levine [1993] p. 553.

⁶ By phenomenal concepts I mean the concepts of conscious experiences understood qualitatively, in terms of what it is like to undergo such experiences.

⁷ That intuition is commonly referred to as the intuition of absent and inverted qualia, respectively.

say, we can still coherently ask why c-fibers firing should feel the way pain does or, indeed, why it should feel any way at all. And without having an answer to this question, the identity of pain and c-fibers firing will not be fully intelligible. For without having an answer to this question, it simply won't be intelligible to us why c-fibers firing should feel the way pain does or why it should feel any way at all. The trouble, of course, is that it seems that the only way to make this connection between c-fibers firing and the way pain feels intelligible is by explaining the phenomenal character of pain in terms of c-fibers firing and that, by assumption, cannot be done. This is, then, the sense in which, according to Levine, we can speak of the explanatory gap between the phenomenal and the physical.

It is worth pointing out that Levine's reasoning could be blocked by denying the crucial assumption that Levine accepts, namely that phenomenal concepts are not functional or causal role concepts. Indeed, there are philosophers who argue that consciousness does not seem to be functional only because we do not know enough about it. As our knowledge increases, we will be able to see that it is certain complex functions that make experiences conscious experiences⁸. If so, the explanatory gap will disappear. For once we see consciousness as a functional phenomenon, we will be able to explain it in physical terms. In what follows I will not pursue this line of thought. I will assume, following Levine, that phenomenal concepts are not functional concepts. That view has certainly an intuitive appeal and the opposite view runs into serious difficulties⁹. Secondly, it is not clear whether the explanatory gap arises even assuming that phenomenal concepts are not functional. There are philosophers who argue that the intuition of the explanatory gap is an illusion even assuming that phenomenal concepts are not functional. Paradoxically, some of them argue that the gap is not real precisely because phenomenal concepts are not functional¹⁰. In what follows, I will argue that the gap is real.

⁸ Chalmers discusses this view as a version of what he calls type-C materialism. This view has recently been endorsed by Thomas Nagel. See Nagel [2002].

⁹ See Chalmers [2002].

¹⁰ See the discussion below.

2. Identities need no explaining

Let me explain further why the question of the intelligibility of psychophysical identity arises even though we have all the empirical evidence for the truth of the identity. The best way to do that is by responding to some of the most common objections to the idea that the explanatory gap is real. One of the most common objections is captured in the slogan “identities need no explaining”¹¹. The point of the objection is that once you are justified in believing in the truth of a given identity, there is no further question to be asked as to why the identity should be true. If “two” entities are one, they simply are and there is no need to explain that. Here is one common illustration of this point. Suppose you believed that Samuel Clemens and Mark Twain were different persons and that suddenly you are told they are the same person. This can prompt you to ask about the reasons to believe that they are one person. So you might ask, for example, why this one person should have different names. But it makes no good sense to ask further why Samuel Clemens and Mark Twain *are* identical. The single person in question couldn’t possibly have been two people.

Here is another illustration of the point that identities need no explaining. Think about the identity of water and H₂O. We can certainly explain why H₂O plays the causal role that we associate with water: why it freezes and boils at such and such temperatures, why it is liquid at room temperature, etc. But we cannot explain why H₂O is water. For explaining that would be explaining why H₂O is itself and that makes no sense.

Now, while I agree that identities as such need no explaining, I do not think that this dissolves the problem of the explanatory gap. Think about the “gap” question again. Why should the physical property *P* feel like that? If this question is interpreted as the request for the explanation of the identity itself, then this question makes no sense and there is no explanatory gap. However, it is far from clear that this is how this question should be interpreted. The question seems to make perfect sense and if it loses its sense when it is interpreted as asking about

¹¹ The slogan comes from Block and Stalnaker [1999].

the explanation of the identity, the conclusion to draw is that this is simply not how the question should be understood. The key task, of course, is to see what sense can be made of the “gap” question. My proposal is given in the next section.

3. The “gap” question properly understood

We seem to think that the “gap” question is the question asking about the explanation of psychophysical identity once we assume that this is not the question about the justification of psychophysical identity, that is, once we assume that the question arises even if we assume that we have all the empirical evidence for believing that the property *P* and a given sort of feeling are identical. But now, if the “gap” question does not make sense as the question about the justification of psychophysical identity and if it does not make sense as the question about the explanation of the identity either, then what exactly is the sense of this question, assuming that the question does make sense?

The answer to this question depends on distinguishing *two concept-dependent readings* of our “gap” question. I take it that the “gap” question cannot make sense on the purely referential reading since on this reading it would amount to the question about the explanation of identity and, as we have just seen, identities need no explaining. The “gap” question can make sense only on the concept-dependent reading. But now, assuming that the “gap” question makes sense on the concept-dependent reading and that it is not a question about the justification of the identity between *P* and a given kind of feeling, the difficulty is to see why it should be any different from the question about justification. On one natural concept-dependent reading, our “gap” question may be understood to ask why the physical concept expressed by “*P*” and the phenomenal concept expressed by “this feeling” have the same referent and, clearly, on this reading the question would amount to the question about justification¹². It seems clear to me,

¹² This is how the “gap” question is interpreted by most of the commentators. See, for example, Papineau [1999] and Tye [1999]. Indeed, Papineau argues that we need not be bothered by the “gap” question, interpreted as the question about the justification of psychophysical identity, since there is a perfectly good answer to it, namely one that appeals to the fact that phenomenal and physical states always appear at the same place in the causal scheme of things. Paradoxically, this is precisely the answer that Levine assumes is not the answer to his “gap” question. But this, of course, brings us back to the question of how the “gap” question should be understood.

though, that there is another concept-dependent reading of our “gap” question. The question may be understood to ask why the physical property *P* should be *conceived of* under the description “feels like this”. This is a perfectly natural question to ask and yet it is different from the first concept-dependent question. It is one thing to ask why the physical property *P* should be the referent of the concept expressed by “feels like this” and another to ask why *P* should be conceived of under the *phenomenal description* associated with the phenomenal concept in question. In order to answer the first question one only needs to point out that *P* plays the relevant causal role that we associate with a given kind of feeling. But knowing that surely does not suffice to answer to the second question. For even if we know that *P* plays the relevant causal role, we can still coherently ask why *P* should be *described* in terms of this feeling or why it should be described in terms of any feeling at all. Intuitively, without having an answer to this question, the identity of *P* and this feeling will not be fully intelligible.

Unfortunately, I do not think there is an answer to our question. If phenomenal character were explicable in physical terms, the description “feels like this” would be entailed a priori by the relevant physical description and then there would be a straightforward answer to our “gap” question. For then we might say that *P* feels like this simply in virtue of being the kind of physical state it is. An analogy may be helpful here. Consider the identity of water and H₂O. Given that this identity is true, there are two different descriptions of water, the description in terms of watery stuff and the description in terms of H₂O. However, the latter description entails the former and this explains why water (H₂O) should be described as watery stuff. Since there is no corresponding entailment between physical and phenomenal descriptions, it becomes utterly mysterious why any relevant physical state should be described phenomenally. There simply seems to be no answer to this question and in this sense, I claim, psychophysical identity is not intelligible. Assuming that phenomenal states are physical, it is a brute and inexplicable fact about certain physical states that they should be described phenomenally.

4. One property under different descriptions

I have argued that assuming that phenomenal descriptions are not entailed a priori by physical descriptions, there is no explanation of why physical states should be described phenomenally. Here is another way of looking at this difficulty. Assuming that phenomenal descriptions are not entailed a priori by physical descriptions, phenomenal descriptions are not physical descriptions. But then it remains unclear why physical states should have non-physical descriptions. That just seems incoherent.

Assuming that phenomenal descriptions are not physical descriptions and assuming that phenomenal and physical descriptions are true of the same properties, it seems that you cannot say about the properties they are true of that they are physical in a sense that is not relativized to the way they are described. The proper thing to say would be that those properties are physical relative to their physical description and that they are phenomenal, that is, not physical, relative to their phenomenal description. But that would imply, of course, that the nature of the properties in question is quite obscure.

Physicalists would certainly reply that this line of thinking begs the question. They might say that it is not incoherent to say that phenomenal properties are physical even relative to their phenomenal description. After all, they might say, even when we consider phenomenal properties relative to their phenomenal description, we assume that they are identical with physical properties. But here is the key point: by assumption, the physical properties that phenomenal properties are supposed to be identical with are themselves described phenomenally, that is, non-physically. If so, we might say that they are physical only relative to their physical description and the physicalist response carries no weight.

5. The antipathetic fallacy

The source of the explanatory gap has never been well understood. For one thing, it has not been clear why there should be any troublesome explanatory gap given that we have all the empirical evidence for the truth of psychophysical identity. I

believe that the account of the explanatory gap I have suggested here makes this matter more clear.

There is another account of the above paradox that is discussed in the literature, so let me briefly contrast it with my account. This other account appeals to the fact that the deployment of phenomenal concepts is apt to trigger in us the images of the very experiences that phenomenal concepts refer to. Take the phenomenal concept *red*. When we exercise this concept, it is apt to trigger in us a visual image of red. Similarly, the concept of pain is apt to trigger in us an image of pain. An image of an experience and the imagined experience are not identical but they are obviously phenomenally akin. What it is like to imagine seeing something red corresponds to what it is like to see something red. Similarly, an imagined pain shares some of the phenomenal unpleasantness of a real pain. We might say then that when we use our phenomenal concepts imaginatively we bring to mind, in a literal sense, an instance of the experiential property we are thinking about. Nothing like that happens, of course, when we use physical concepts. Those concepts are not apt to trigger in us any images of experiences they putatively refer to and the suggestion is that it is in this sense that physical concepts *leave out* the phenomenology of experiences. Indeed, since the phenomenology of experiences is left out, it might seem that physical concepts cannot refer to experiences.

The common response to this intuition of the distinctness of phenomenal and physical properties is that it is an illusion. Indeed, Papineau [1999] refers to this intuition as the “antipathetic fallacy”. From the fact that physical concepts do not bring to mind, in the sense explained above, the experiences that they are supposed to stand for it simply does not follow that they do not refer to those experiences. It follows that there is a difference in the way physical and phenomenal concepts function but not that they have different referents.

It should be clear that my account of the explanatory gap is different since I do not locate the crucial difference between phenomenal and physical concepts at the level of their functioning but rather at the level of the modes of presentations associated with those concepts. Assuming that phenomenal concepts refer to

physical properties, those concepts do not represent physical properties as physical. Instead, they represent them as phenomenal properties, under the mode of presentation “this is what it is like”. This is then what generates the explanatory gap, on my view. The question that requires answering is why physical properties should be described phenomenally.

Now, I have argued that this question cannot be answered by assuming that we have reasons to believe that psychophysical identity is true. Thus, the response to the intuition of the explanatory gap described above will not work as a response to our question. In response to our question it wouldn't help to say that from the fact that the modes of presentation associated with phenomenal and physical concepts are different it does not follow that those concepts do not have the same referents. We can grant that we have reasons to think that the referents of phenomenal and physical concepts are the same. The problem is that this falls short of making it intelligible why physical properties should be described phenomenally.

6. Phenomenal character needs no explaining: first case

I have identified one plausible source of the intuition of the explanatory gap and I have defended my account against some arguments designed to show that the explanatory gap is unreal. In the remaining part of the paper, I will defend my account against some further objections of this sort.

On an intuitive understanding of the explanatory gap, the gap arises because we think of psychophysical identity that while it cannot be explanatory it *should be*. Thus, one reason why we might deny that the explanatory gap is real is that we might deny that there are any reasons why we should expect psychophysical identity to be explanatory. There are influential philosophers who think that this is precisely the case. They argue that the expectation in question is simply an illusion. Broadly speaking, there are two ways in which this point is made. One way is to argue that we should not expect psychophysical identity to be explanatory due to the nature of phenomenal concepts. The other way is to argue that this expectation is an illusion since the intuitions that generate it are

illusory. For accounts of the second sort, the nature of phenomenal concepts is crucial, too, since on accounts of this sort it is the very nature of phenomenal concepts which explains why the intuitions that are supposed to support our expectations with respect to psychophysical identity are illusory.

I think it is quite obvious that the first line of argument must fail once we have seen that there is a good conceptual reason to expect psychophysical identity to be explanatory. I also think that the second line of argument fails since the intuitions that are taken to generate the intuition of the explanatory gap and that are assumed to be illusory are different from the intuition that I have accounted for here. Thus, even if we grant that the other intuitions are illusory, we are free to assume that the intuition of the gap is generated by the intuition that I have accounted for here.

It is easier to see why the first line of argument fails, so let me take up the first line of argument first. This is the line developed by Michael Tye [1999]. According to Tye, psychophysical identity need not be fully explanatory, that is, it need not provide the explanation of phenomenal character since it is guaranteed by the nature of phenomenal concepts that phenomenal character cannot be explained. Phenomenal concepts are not causal role concepts and only causal roles can be explained in physical terms. Thus, it is guaranteed by the nature of phenomenal concepts that phenomenal character cannot be explained. Now, why should that imply that phenomenal character needs no explaining? Tye's reasoning is as simple as that. Something needs explaining only if it can in principle be explained. So since it is a conceptual truth that phenomenal character cannot be explained, phenomenal character cannot in principle be explained and hence there is no issue of how to explain it. Here is how Tye summarizes the point of his response to the problem of the explanatory gap:

The so-called "explanatory gap" derives largely from a failure to recognize the special features of phenomenal *concepts*. These concepts, I maintain, have a character that not only explains why we have the intuition that something

important is left out by the physical (and/or functional) story but also explains why this intuition is not to be trusted¹³.

I am not convinced by Tye's reasoning¹⁴. The move from the premise that phenomenal character cannot be explained due to the nature of phenomenal concepts to the conclusion that phenomenal character needs no explaining is too quick. This is because we have a plausible story to tell us why due to the very nature of phenomenal concepts we feel that phenomenal character does need explaining. This is precisely the story that I gave earlier. It certainly sounds paradoxical that reflection upon the nature of phenomenal concepts makes us expect psychophysical identity to be explanatory and that at the same time the nature of those concepts guarantees that our expectation cannot be fulfilled. But the mere fact that our expectation cannot be fulfilled does not show that our expectation is illusory so long as the expectation is supported by plausible conceptual reasons. To show that the expectation is illusory one would need to show in addition that the conceptual reasons that we think ground our expectation—whatever they are—are not, in fact, good reasons. That is precisely what the second line of argument against the explanatory gap attempts to achieve.

Tye himself does not see any *prima facie* good motivation for the expectation that phenomenal character should be explicable in physical terms and this may be another reason why he feels so confident in assuming that phenomenal character simply needs no explaining. That he does not see any such motivation is evidenced by the fact that he allows only two *prima facie* plausible interpretations of the question "Why does physical state *P* feels like this?": the referential interpretation and one according to which the "gap" question is understood as asking for the justification of why *P* and the relevant phenomenal state are identical. Obviously, on neither of those interpretation is there any need to explain phenomenal character in physical terms. As we saw, on the referential interpretation the "gap" question does not really make sense since on this

¹³ Tye [1999] p. 707.

¹⁴ A similar line of argument is taken by Papineau [1999].

interpretation the question must be understood as asking for the explanation of the identity itself and, as we saw, identities need no explaining. On the other hand, when we interpret the “gap” question as the question about the justification of phenomenal-physical identity, there is no need to explain phenomenal character either, since the justification of the identity can be based solely on the explanation of the causal role of phenomenal states¹⁵.

From what I have said so far, however, it should be clear that the “gap” question cannot be understood as the question about the justification of psychophysical identity. The “gap” question arises even if we assume that the identity is justified.

7. Phenomenal character needs no explaining: second case

Let me now turn to the second line of argument against assuming that we should expect psychophysical identity to be explanatory. This is the line that is taken by Brian Loar [1997, 1999]. Loar’s argument depends on assuming that there are some articulate reasons why we might expect psychophysical identity to be explanatory. The point of the argument is then that those are not good reasons, that is, the point is that we are under an illusion when we think that those reasons justify the expectation that psychophysical identity should be explanatory.

Loar’s response to the problem of the explanatory gap comes, in fact, in two steps, so let us proceed in accordance with those steps. First, and indeed quite surprisingly, Loar assumes that there is no mystery as to how phenomenal concepts can pick out physical properties, say, the properties of the brain. Phenomenal concepts belong to the wider class of *recognitional concepts* and they pick out physical properties in the manner of all other recognitional concepts. What makes a given concept a recognitional concept? Roughly, a recognitional concept for Loar is a type-demonstrative of the form “x is one of *that* kind”, where x can be a physical thing of a certain perceived kind or an internal state accessible

¹⁵ Tye assumes that there is a whole range of facts about conscious experiences that go beyond their causal role and that can be explained in physical terms. Still, phenomenal character itself remains unexplained on Tye’s account.

experientially through introspection. The possession of such concepts is grounded in dispositions to classify objects, events and situations in the way of perceptual or experiential discriminations. Thus, the possession of such concepts does not require any theoretical background, instead it requires that one has actually had the experiences that fix the reference of those concepts or that one has the disposition to recognize those experiences if one has them. An example of a recognitional concept which is not a phenomenal concept is the concept of cramp. This is a recognitional concept since it picks out its referent (muscle contraction) by discriminating it experientially, under the experiential mode of presentation “that sort of state”¹⁶.

Now, although the concept of cramp picks out its referent under an experiential (phenomenal) mode of presentation, that concept is not a phenomenal concept since it does not pick out a phenomenal property, that is, it does not pick out the property of its being like this to have cramp feelings; that phenomenal property only fixes the reference of that concept. Things are different in the case of phenomenal concepts. The phenomenal properties that fix the reference of phenomenal concepts are identical with the properties these concepts pick out and, in this sense, phenomenal concepts pick out their referents *directly*¹⁷.

If there is this crucial difference between phenomenal concepts and non-phenomenal recognitional concepts, then Loar’s claim that phenomenal concepts pick out physical properties in the manner of all recognitional concepts does not seem to clarify what it purports to clarify after all. It is easy to see how it is that the

¹⁶ Although recognitional concepts may somehow incorporate or display a particular perception or a quale-token, they are not, on Loar’s account, definite descriptions of the form “the kind of state that causes this perception” (nor are they descriptions of the form “the kind of state that is tracked by this identifying disposition” or “the property I hereby ostend”, etc.). To regard such concepts as definite descriptions would be psychologically unmotivated, according to Loar. Thus, the concept of cramp is not the definite description “the kind of state that causes this feeling”. This means that the reference of the concept of cramp and all other recognitional concepts is fixed causally rather than via descriptions. For more on recognitional concepts, see Loar [1997] and Loar [2003].

¹⁷ As Loar points out, it is tempting to say that phenomenal modes of presentation, the modes of presentation of the form “that is what it is like”, present the referents of phenomenal concepts as they are essentially. But Loar denies that this is the case. Phenomenal concepts do not reveal the essence of phenomenal properties. For assuming that the essence is physical, the essence cannot be grasped a priori, simply in virtue of understanding phenomenal concepts and the relevant physical concepts.

concept of cramp picks out muscle contraction: the reference of the concept of cramp is fixed by cramp feelings and the concept of cramp picks out muscle contraction since muscle contraction causes cramp feelings. But since phenomenal concepts refer directly, the feelings that fix the reference of those concepts cannot be caused by the physical properties that those concepts are supposed to pick out. The feelings in question have to be identical with the physical referents of those concepts. This raises a difficulty. For it now appears that we should be able to see a priori, simply in virtue of understanding phenomenal concepts and the relevant theoretical-physical concepts, that they pick out the same properties, and that simply sounds implausible.

Loar refers to the above expectation as *the expectation of transparency*. Indeed, he argues that this is precisely the expectation that generates the intuition of the explanatory gap. For the expectation that we should be able to see a priori that phenomenal and physical concepts pick out the same properties is the expectation to the effect that physical concepts should entail a priori phenomenal concepts and hence that phenomenal properties should be explicable in terms of physical properties. The same sort of expectation arises in the case of directly referring non-phenomenal concepts, such as the concept of liquidity. Indeed, in the case of those other concepts, the expectation of transparency is fulfilled. Thus, the concept of liquidity is entailed a priori by the relevant physical concept since the theory of liquids entails a priori that the functional description that we associate with the concept of liquidity is realized. The point to be emphasized is that the concept of liquidity refers directly: the functional description that we associate with the concept of liquidity is not the description of any properties distinct from the property of being liquid. This is precisely why we would expect this description to be entailed a priori by the relevant physical theory. And similarly, since phenomenal concepts refer directly, we expect that the properties they pick out should be deducible from the relevant physical theory as well. Unfortunately, in the case of phenomenal concepts, this expectation cannot be fulfilled.

Now, Loar argues that if this is what generates the explanatory gap, the gap is unreal. This is because the expectation of transparency is an illusion, according

to Loar. It is simply an illusion to expect that phenomenal concepts should be entailed a priori by physical concepts. We could expect there to be such an entailment if phenomenal concepts were functional concepts. But phenomenal concepts are not functional but recognitional concepts and as such they are conceptually independent of physical concepts. This is the second step of Loar's response to the problem of the explanatory gap. The key idea here is that there is no inconsistency between assuming that phenomenal concepts refer to physical properties directly and seeing phenomenal concepts as conceptually independent of physical concepts. This lack of inconsistency is guaranteed by Loar's analysis of phenomenal concepts as recognitional concepts. Assuming that phenomenal concepts are recognitional concepts, their conceptual independence of physical concepts can be seen as a matter of the difference between the conceptual roles of those two sorts of concepts: by assumption, phenomenal concepts discriminate physical properties without analyzing them in scientific terms. This conceptual difference is not affected by the fact that phenomenal concepts refer directly. Thus, it is hardly surprising that phenomenal concepts cannot be entailed by physical concepts even if we grant that phenomenal concepts refer to physical properties directly. Here is how Loar summarizes the point of his response to the problem of the explanatory gap:

The problem of the explanatory gap stems from an illusion. What generates the problem is not appreciating the fact that there can be two conceptually independent "direct grasps" of a single essence, that is, grasping it demonstratively by experiencing it, and grasping it in theoretical terms. The illusion is of *expected transparency*: a direct grasp of a property ought to reveal how it is internally constituted, and if it is not revealed as physically constituted, then it is not so. The mistake is the thought that a direct grasp ought to be a transparent grasp, and it is a natural enough expectation¹⁸.

Here is my response to Loar. I do not think that Loar dissolves the problem of the explanatory gap. This is because I do not think that what Loar has identified as the

¹⁸ Loar [1997] p. 609.

source of the intuition of the explanatory gap is the only possible source of that intuition. To repeat, on Loar's account, the intuition stems from the expectation of transparency and the expectation of transparency stems from the intuition that, given that phenomenal concepts refer directly, *no empirical investigation should be required* in order to see that phenomenal and physical concepts pick out the same properties. However, on the account I proposed earlier, it is natural to think that the source of the explanatory gap is different. On my account, the intuition of the explanatory gap stems from the intuition that *no empirical investigation is sufficient* to make psychophysical identity fully intelligible. Empirical investigation into the nature of phenomenal properties can only lead us to the explanation of the causal role of those properties and while that sort of explanation gives us sufficient justification to believe that psychophysical identity is true, it falls short of making the identity fully intelligible for the reasons I have given earlier. This, on my account, is the reason why we expect that we should be able to explain not only the causal role of experiences but also their phenomenal character. The explanation of the latter seems to be required in order to make psychophysical identity fully intelligible. I do not think that anything Loar has said makes this intuition illusory.

8 Conclusion

I have argued that there is a very natural intuition which explains why the feeling that there is the troublesome explanatory gap seems so strong. The intuition is that assuming that phenomenal character cannot be explained in physical terms, we do not understand why physical states should be described phenomenally and, in this sense, psychophysical identity remains unintelligible. Some influential accounts of the problem of the explanatory gap attempt to dissolve this problem but, as I have shown, those accounts fail since they fail to capture the intuition that really generates the problem.

References

- Block and Stalnaker [1999] – N. Block and R. Stalnaker, *Conceptual Analysis, Dualism, and the Explanatory Gap*, “The Philosophical Review” (108) 1999, p. 1-46.
- Chalmers [1996] – D. Chalmers, *The Conscious Mind*, Oxford University Press, Oxford 1996.
- Chalmers [2002] – D. Chalmers, *Consciousness and Its Place in Nature*, in: *Philosophy of Mind*, ed. D. Chalmers, Oxford University Press, Oxford 2002.
- Gendler and Hawthorne [2002] – T. S. Gendler and J. Hawthorne (eds.), *Conceivability and Possibility*, Oxford University Press, Oxford 2002.
- Levine [1983] – J. Levine, *Materialism and Qualia: The Explanatory Gap*, “Pacific Philosophical Quarterly” (64) 1983, p. 354-61.
- Levine [1993] – J. Levine, *On Leaving Out What It Is Like*, in: *Consciousness*, ed. M. Davies and G. Humphreys, Blackwell, Oxford 1993, p. 543-555.
- Loar [1997] – B. Loar, *Phenomenal states*, in: *The Nature of Consciousness*, ed. N. Block, The MIT Press, Cambridge, Mass. 1997, p. 597-616.
- Loar [1999] – B. Loar, *David Chalmers’s “The Conscious Mind”*, “Philosophy and Phenomenological Research” (59) 1999, p. 465-72.
- Loar [2003] – B. Loar, *Qualia, Properties, Modality*, “Philosophical Issues” (13) 2003, p. 113-129.
- McGinn [1989] – C. McGinn, *Can we solve the mind-body problem?*, “Mind” (98) 1989, p. 349-66.
- Nagel [2002] – T. Nagel, *Psychophysical Nexus*, in: *Concealment and Exposure*, ed. T. Nagel, Oxford University Press, Oxford 2002, p. 194-235.
- Papineau [1999] – D. Papineau, *Mind the Gap*, “Philosophical Perspectives” (12) 1999.
- Tye [1999] – M. Tye, *Phenomenal Consciousness: the Explanatory Gap as a Cognitive Illusion*, “Mind” (108) 1999, p. 705-25.