

*Can phenomenal concepts explain the explanatory gap?*¹

ESA DÍAZ-LEÓN

University of Sheffield

Pip02ed@sheffield.ac.uk

An important argument against physicalism is the so-called *conceivability argument*. Intuitively, this argument claims that since certain statements concerning the separation of the physical and the phenomenal are conceivable, they are possible. This inference from conceivability to possibility has been challenged in numerous ways. One of these ways is the so-called *phenomenal concept strategy*², which has become one of the main strategies against conceivability arguments. David Chalmers says it “is perhaps the most attractive option for a physicalist to take in responding to the problem of consciousness”.³ Certainly, in the recent years, a multitude of proposals of that sort have been proposed and developed.⁴ However, Chalmers (2006) has recently argued that no version of the phenomenal concept strategy can succeed. In what follows, I will examine his main argument for that conclusion, and I will argue that it is not sound. I will conclude that he has not posed any serious problem for the phenomenal concept strategy to succeed in blocking conceivability arguments.

1. Conceivability arguments and the phenomenal concept strategy

Conceivability arguments can be formulated simply thus: let P be a complete physical description of the actual world. Let Q be a complete phenomenal description of the world. It seems that we can imagine a scenario where P holds but Q does not. For example, we can imagine a possible world physically identical to our world that lacks consciousness. That is, $P \rightarrow Q$ is not a priori true. There is an epistemic gap between our knowledge of physical truths and our knowledge of phenomenal truths: knowledge of the former does not suffice for having knowledge of the latter. This can be further motivated by reflecting on Frank Jackson’s Mary, who knows all physical facts, but does not know what it is like to see red.

Physicalism is committed to the claim that the conditional $P \rightarrow Q$ is necessary. But as we have seen, it is a posteriori, that is, it is conceivably false. From this, some argue that the conditional is not necessary. So physicalism is false.

What motivates this inference from the conceivability of $P \& \sim Q$ to its possibility? There are different (but related) ways of motivating this inference. A general way of putting it is the following. If the conditional $P \rightarrow Q$ was necessary, it would be an a posteriori necessity (since we have granted it is not a priori). But, the argument goes, all standard cases of a posteriori necessities satisfy a certain model, namely, it is always the case that when we can conceive of a statement being false, there is a

¹ I am grateful to the AHRC and the University of Sheffield for financial support. I also wish to thank Stephen Laurence for helpful comments on an earlier version of this paper.

² This label is first introduced by Daniel Stoljar 2005.

³ Chalmers 2006, 1.

⁴ See, for instance, Loar 1997, 1999 and 2003, Hill 1997 and Hill & McLaughlin 1999, Papineau 2002, Sturgeon 2001.

possible world that *verifies* that statement.⁵ For example, consider the following sentences:

- (1) Water = H₂O
- (2) Heat = Molecular motion

We can easily conceive of these sentences being false. In each case, there is a corresponding scenario where what we imagine, or something suitably related, is true. For the negation of (1), the verifying possible world could be the Twin Earth, where watery stuff is not H₂O. For the negation of (2), we can think of a possible world where the phenomenon that causes sensations of heat is other than molecular motion.

These examples motivate the thesis that the conceivability of any statement must be accompanied by the possibility of either that very same statement, or some other closely connected, being true. Therefore, the conceivability of $P \& \sim Q$ must be accompanied by some relevant possibility. And in this case, the relevant possibility is none other than the possibility of a world where P holds but Q does not. So physicalism must be false.

This is the conceivability argument, in a nutshell. The phenomenal concept strategy attacks it in the following way. As we have seen, there is an epistemic gap between physical truths and phenomenal truths: P does not entail Q a priori. Some argue this implies the existence of an ontological gap between physical truths and phenomenal truths. But the phenomenal concept strategy claims that the epistemic gap is not due to any ontological gap, but rather to the special features of phenomenal concepts.

The task of the phenomenal concept strategy is to provide an alternative explanation of why the conditional $P \rightarrow Q$ is a posteriori. This new explanation does not rely on the conditional being either contingent or false. Rather, the explanation focuses on some features of phenomenal concepts and their different character from that of physical concepts. Then, this explanation is *compatible* with both the conditional being necessary and the conditional being contingent: since phenomenal concepts are as they are, $P \rightarrow Q$ has to be a posteriori, independently of its modal status. If this explanation of the a posteriori character of the conditional is a plausible alternative, then the conditional's being a posteriori does not entail that it is not necessary. For being a posteriori is compatible with being necessary.

Chalmers claims that any successful version of the phenomenal concept strategy should have the following structure:

Proponents put forward a thesis C attributing certain psychological features –call these the key features– to human beings. They argue (i) that C is true, i.e. that humans actually have the key features; (ii) that C explains our epistemic situation with regard to consciousness, i.e. that C explains why we are confronted with the relevant distinctive epistemic gaps; and (iii) that C itself can be explained in physical terms, i.e. that one can (at least in principle) give a materialistically acceptable explanation of how it is that humans have the key features. (*Op. cit.*, 5)

⁵ A possible world verifies a statement iff the statement is true at that possible world considered as *actual*. I will try to keep the discussion free from the technicalities of two-dimensional semantics, where possible.

Chalmers argues that while the three elements (i) to (iii) are essential for the strategy to work, at the same time no account of phenomenal concepts can satisfy both (ii) and (iii). Therefore, he concludes that any version of the phenomenal concept strategy is irredeemably condemned to fail in resisting the inference from the epistemic gap to an ontological gap. In the next section, I will explain Chalmers' argument for this claim. In the last section, I will argue that his argument does not work.

2. *A Dilemma for the Phenomenal Concept Strategy*

The main problem of the phenomenal concept strategy, according to Chalmers, is that it cannot simultaneously satisfy requirements (ii) and (iii). For any account, if it satisfies (ii) it cannot satisfy (iii), and *vice versa*. His argument for this claim is the following:

- (1) If $P \& \sim C$ is conceivable, then C is not physically explicable.
 - (2) If $P \& \sim C$ is not conceivable, then C cannot explain our epistemic situation.
-
- (3) Either C is not physically explicable, or C cannot explain our epistemic situation.⁶

The conclusion asserts that for every account of phenomenal concepts C , either (ii) is not satisfied (that is, C cannot explain our epistemic situation), or (iii) is not satisfied (that is, C cannot be explained in physical terms). I will now explain each premise in turn.

This argument has the form of a dilemma. The dilemma starts in the following way: for any account C , either it is entailed a priori by physical truths (P) or it is not. That is, either $P \& \sim C$ is conceivable, or it is not. In order to see whether $P \& \sim C$ is conceivable, we have to try to imagine some physical duplicate of the world which does not satisfy C . In particular, we can focus on zombie-worlds: physically identical duplicates of this world where no-one is conscious. According to the proponent of the phenomenal concept strategy, zombie-worlds are conceivable. Zombie-worlds clearly satisfy P . Then we can ask ourselves whether zombie-worlds would also satisfy C . If they do not, then $P \& \sim C$ is conceivable. What happens if $P \& \sim C$ is conceivable? Well, the problem is that, in this case, P does not entail C a priori, and then, according to Chalmers' characterization of explanation as a priori entailment, we cannot explain C in physical terms.

Chalmers has argued that in order to explain some macro-phenomenon in microphysical terms, we have to show how microphysical facts give rise to that macro-phenomenon, and this requires that microphysical facts transparently show how the macro-phenomenon obtains. If microphysical facts did not entail a priori the instantiation of the macro-phenomenon, then we would require some additional explanatory materials in order to explain such an instantiation, and this would show that the original explanation was incomplete.

According to Chalmers, in order for the phenomenal concept strategy to work, it has to show that physicalism is compatible with the epistemic gaps, and therefore, its proponents have to show how physical facts alone give rise to the key features of phenomenal concepts which in turn give rise to the relevant epistemic gaps. If they do not provide a physically acceptable explanation of the key features of phenomenal

⁶ Chalmers, *Op. Cit.*, 7.

concepts, then they will not have motivated the view that their explanation of the epistemic gap is indeed compatible with physicalism.

The second premise of the argument examines the possibility that $P \& \sim C$ is not conceivable, that is, that P does entail C a priori. In this case, even zombie-worlds satisfy C : every physical duplicate of the actual world satisfies C . This horn of the dilemma represents the possibility that we are really capable of offering a physically acceptable explanation of the key features of phenomenal concepts. Then, we would not have the problem of the previous case, but we would have another problem: now even zombies satisfy C . That is, this account of phenomenal concepts does not discriminate between zombies and conscious people. Arguably, zombies have a very different epistemic position than us. We are conscious, but they are not. We have beliefs about our own phenomenal states, but they do not. We have knowledge of phenomenal facts, but they do not. Our epistemic positions are as different as they might be. Our epistemic position (let's call it E), which concerns the truth-values of our beliefs and the epistemic status of our beliefs (that is, justified or unjustified) is very different from zombies' epistemic position. That is, we satisfy E , whereas zombies do not satisfy E . So zombies satisfy C (according to this horn of the dilemma), but not E . Then, C does not entail E a priori, because there are beings (zombies) that satisfy C but not E . And therefore, C cannot explain E . But E is our global epistemic position, which includes the epistemic gap between P and Q . Then, C cannot explain the relevant epistemic gap, because having C does not suffice for being in the relevant epistemic position. So premise (2) is true.

The intuitive idea behind this argument is that if some account of phenomenal concepts is tame enough to be explained in physical terms, then it will be satisfied by zombies, but then it cannot explain the relevant epistemic gaps, because zombies' epistemic position differs enormously from ours. They are not conscious, so for them there is no epistemic gap between physical truths and phenomenal truths. But if C cannot explain why the key features of phenomenal concepts give rise to the relevant epistemic gaps, then the phenomenal concept strategy cannot work.

The conclusion is that no account C can both be physically explained and successfully explain the epistemic gap. For if C is a priori entailed by P , this means that zombies will satisfy C , but since zombies do not satisfy E , C cannot explain E . And if C does entail a priori, and therefore can explain, E , then it means that zombies do not satisfy C , and this is because C is not a priori entailed by P . Then, C cannot be physically explained.

In the following section I will argue that the second premise does not hold, that is, it is possible for some account C to be both physically explicable and explain the relevant epistemic gaps.

3. *Explaining the epistemic gap*

In this section I want to raise some problems concerning the second horn of the dilemma, that is, premise (2) above. As we have seen, Chalmers argues that if P entails a priori C , then there will be a transparent explanation of C in physical terms, but the cost is that this account of phenomenal concepts C cannot really explain why there is an epistemic gap between P and Q . The reason is that having C does not entail

such a gap, since zombies satisfy C but are in a different epistemic situation, and therefore do not instantiate the same epistemic gaps as we do.

As I have pointed out in the previous section, Chalmers characterizes someone's epistemic situation as including "the truth values of their beliefs and the epistemic status of their beliefs (as justified or unjustified)" (Chalmers 2006: 9). He assumes we can draw a correspondence between our beliefs and zombies' beliefs: "We assume an intuitive notion of correspondence between the beliefs of a conscious being and the beliefs (if any) of its zombie twin. For example, corresponding utterances by a conscious being and its zombie twin will express corresponding beliefs" (*Ibid*). Even if these utterances do not have the same content, we can still talk about *corresponding* beliefs.⁷ Then, he adds: "We can say that a zombie shares this epistemic situation when the zombie has corresponding beliefs all of which have corresponding truth-value and epistemic status" (9). So in order to evaluate whether zombies share our epistemic situation we have to compare the truth-values and epistemic status of our beliefs and their corresponding beliefs. If there is any difference, our epistemic positions differ. Then, it seems clear enough that zombies do not share our epistemic situation, since zombies' beliefs such as 'I am phenomenally conscious' do not seem to be true of them or at least are not as justified as our corresponding beliefs.

We can agree with this idea, but it is not clear to me why this is a problem for the phenomenal concept strategy. What an account of phenomenal concepts C has to explain is why there is an epistemic gap between P and Q, not our whole epistemic situation. Even if having C does not rule out that one is a zombie, this is not a problem for explaining why P does not entail Q a priori. Zombies do not pose any problem for this explanatory task, because *zombies are not able to infer Q from P either*. This is due to the fact that they cannot understand Q, since arguably they do not possess the concepts required to entertain Q, namely, phenomenal concepts. They do not have phenomenal states, so they cannot have phenomenal concepts (with the same content as ours). Chalmers seems to agree with this point: "It is plausible that since a zombie is not conscious, it cannot have beliefs with exactly the same content as our beliefs about consciousness" (*Ibid*). They can have corresponding beliefs, but with *different* content. Then, they cannot understand our sentence Q, so they cannot deduce a priori our sentence Q from P.

Someone could reply that what is at issue here is rather whether zombies would be able to infer a priori their *corresponding* belief Q from their corresponding belief P. But I do not think that this question is relevant for the phenomenal concept strategy. The reason is this: what C has to explain is why $P \rightarrow Q$, *with the content it actually has*, is a posteriori. And zombies do not pose any problem for that explanatory task, since they are not a counterexample to the following conditional:

(*) If C is true of some beings, then they cannot deduce Q from P a priori

Zombies satisfy C (according to the second horn of the dilemma), but they are not able to deduce Q from P. So nothing here jeopardises the claim that C entails a priori the existence of the relevant epistemic gap between P and Q. Certainly, zombies are

⁷ Maybe these corresponding beliefs would have the same functional role. It seems plausible to say that physical duplicates are also functional duplicates. Then, zombies share our functional description, and then they would have the same functionally characterised beliefs.

not able to infer Q from P for very different reasons than ours: they are not able because they cannot even entertain the propositions involved. This could mean that zombies are irrelevant in order to evaluate whether all beings that satisfy C instantiate the epistemic gap, that is, in order to evaluate (*). But it does not mean that zombies pose a counterexample to (*).

In any case, it is not clear either whether zombies are able to infer a priori their *corresponding* belief Q from their corresponding belief P. This is because it is not clear what the contents of their corresponding phenomenal beliefs are. Kati Balog (1999) suggests that zombies' phenomenal terms will refer to brain states, but Chalmers rejects this view. He argues that Zombie Eliminativist, the one who utters sentences such as "There is no phenomenal consciousness", would be right in a zombie world.⁸ Then, the zombie's phrase "phenomenal consciousness" cannot refer to brain states, because there are plenty of brain states at the zombie world. If zombies' phenomenal terms referred to brain states, then it would be plausible to say that zombies can infer (their corresponding belief) Q from P. But this reply is not open to Chalmers, since he rejects that view. According to Chalmers, the content of zombies' phenomenal beliefs is such that zombie-eliminativists are right when they say that they do not have consciousness. So the content of their beliefs must be somehow related to phenomenal consciousness as we understand it. But if that is true then it is unlikely that zombies can infer a priori Q from P. So they still instantiate the relevant epistemic gap.

Chalmers anticipates this kind of response:

This version concedes that the key psychological features in question cannot explain our full epistemic situation as defined above, but asserts that the features can at least explain certain *aspects* of our epistemic situation, in a way that promises to explain away certain crucial portions of the explanatory gap, or of the other epistemic gaps. [...] A proponent might suggest [...] that such an account can explain the *inferential disconnection* between our physical and phenomenal beliefs, including the fact that the latter are not deducible from the former. (Chalmers 2006, 13)

This response corresponds quite well with what I have been arguing so far, namely, that the phenomenal concept strategy does not have to explain our whole epistemic situation but just why there is an epistemic gap between physical truths and phenomenal truths. And Chalmers seems to accept that this inferential disconnection is also present in the zombies' case, as he claims that "it is plausible that a zombie's physical and quasi-phenomenal beliefs are no more inferentially connected than a conscious being's beliefs" (14). However, he is not satisfied with this response, for the following reasons:

While the inferential disconnection strategy might physically explain an inferential disconnection between physical and phenomenal *beliefs*, the anti-physicalist's crucial epistemic gap involves a disconnection between physical and phenomenal *knowledge*. (*Ibid*)

What Chalmers is claiming at this point is that the question that the phenomenal concept strategy has to explain is why there is a disconnection between our *knowledge* of physical facts and our *knowledge* of phenomenal facts:

⁸ Chalmers 2006, 9-10.

The relevant epistemic gap is characterised in a way that truth and knowledge are essential. [...] It is crucial to the knowledge argument that Mary gains new factual *knowledge*. [...] It is crucial to the conceivability argument that one can conceive beings that lack phenomenal states that one actually has. And it is crucial to the explanatory gap that one has substantive knowledge of the states that we cannot explain". (*Ibid*)

So according to Chalmers, the strategy has to explain how we can know all physical facts without thereby knowing all phenomenal facts (which *we do* know otherwise), and how we can conceive of beings that are physically identical to us but lack the phenomenal states of which *we* have substantive knowledge. The question at issue, then, is the following: what is required in order to provide a satisfactory explanation of the explanatory gap so understood? This is the question I will discuss in the remainder of this section.

It seems clear that it is not enough to merely explain that physical truths do not entail a priori zombies' *pseudo*-phenomenal beliefs. This is related to the option we have considered above. According to this option, explaining the epistemic gap would involve explaining why zombies cannot infer their corresponding phenomenal beliefs from their corresponding physical beliefs. However, Chalmers argues that these epistemic gaps are very different from our substantive epistemic gap, which crucially involves the truth and justification of our phenomenal beliefs. To explain those weakened forms of epistemic gap (that is, why zombies cannot infer their *corresponding* phenomenal beliefs from their corresponding physical beliefs) is not to explain our knowledge-involving epistemic gap.

As I have said before, I think that explaining the epistemic gap requires explaining why physical truths do not entail a priori our phenomenal truths, as we conceive of them, and therefore a putative explanation of why zombies' physical beliefs do not entail their corresponding phenomenal beliefs is not sufficient. However, I do not think that zombies pose any problem for a genuine explanation of the epistemic gap. Let me explain this in some detail. What Chalmers has argued, in the second horn of the dilemma, is that any account of phenomenal concept C that is entailed a priori by P (that is, explained in physically acceptable terms) is going to have the following problem: it will be unable to explain how the key features of phenomenal concepts give rise to the relevant epistemic gap. Chalmers argues for this claim by means of a putative counterexample: if C explained our epistemic situation E, then every being that satisfies C would also satisfy E. However, zombies satisfy C but not E. So C does not explain E.

I have argued in response that this counterexample is not a good one, because zombies satisfy both C and the relevant features of E, namely, they are not able to infer Q a priori from P. In his discussion of this kind of response, Chalmers qualifies the counterexample: zombies satisfy C but do not experience the epistemic gap *as we do*, because they cannot have substantial knowledge of phenomenal truths. Chalmers claims that the strategy has to explain how the key features of phenomenal concepts give rise to our *knowledge-involving* epistemic gap; and he argues that zombies (may) instantiate C but not our knowledge-involving epistemic gap.

I think that his counterexample only works if we equivocate on the expression "knowledge-involving epistemic gap". This is ambiguous, for it can be understood in two ways. There is a *weak* reading (i), according to which the strategy has to explain

something more than just how physical truths fail to entail a priori zombies' *pseudo-phenomenal* beliefs. That is, according to this reading, C should explain why P fails to entail Q a priori, with the content it actually has. There is also a *strong* reading (ii), according to which the strategy has to explain how *we can have* a substantive phenomenal knowledge of phenomenal states and at the same time, we fail to infer it a priori from physical truths. The difference between these two readings is that the second, unlike the first, entails that all subjects that satisfy C have to have a substantive phenomenal knowledge of their genuine phenomenal states (and not merely pseudo-phenomenal), in order for C to provide a successful explanation of the epistemic gap.

I think that the demand for an explanation in the second, stronger sense is not motivated: why should we explain the special features of phenomenal knowledge *only in terms* of the psychological features of phenomenal concepts? The psychological features of phenomenal concepts were not invoked, by the proponents of the strategy, in order to explain the special kind of knowledge that we have when, say, we see red for the first time; rather, these features are invoked in order to explain why this special knowledge cannot be *inferred* from physical knowledge. The second question is different from the first.

My suggestion is that the strategy should aim to explain the explanatory gap only under the weak reading, not under the strong one. But Chalmers' counterexample only works when it is addressed against explanations of the explanatory gap in the stronger sense.

In order to construct a relevant counterexample against reading (i), Chalmers should do the following. He should find a (conceivable) case in which an individual satisfies C but does not instantiate the epistemic gap, for instance, because the individual is able to infer phenomenal knowledge from physical knowledge. But the counterexample that Chalmers has provided is not like this. He has just offered a case in which some individuals (zombies) satisfy C but do not have substantive knowledge of phenomenal states. This is not surprising: they are zombies! What is crucial to my reply to Chalmers is that this counterexample is not a counterexample to the most sensible reading of the explanatory task that the strategy has to accomplish, namely, task (i). For zombies satisfy C, and at the same time, they are not able to *infer* substantive phenomenal knowledge from physical phenomenal knowledge. They cannot infer such phenomenal knowledge because they cannot have such genuine phenomenal knowledge. But this does not matter: the putative counterexample against C entailing E is just not a good counterexample, because it is not a case of C without E.

There is another (closely related) way of reading Chalmers' objection that could seem problematic. He says that it is crucial to the conceivability argument that we can conceive of beings physically identical but lacking phenomenal states *that one actually has*. Otherwise, the conceivability argument "would not get off the ground" (14). We can accept this point, but this does not force us to maintain that the phenomenal concept strategy has to explain how we are able to get conceivability arguments off the ground. That is, I do not think it is fair to ask the strategy to provide an account of phenomenal concepts such that it entails a priori all the necessary conditions for a subject to be able to formulate a conceivability argument. There are

some conditions that no plausible account C could entail, such as the condition that we know some basic logic. But clearly this is not a problem for the strategy. Likewise, it does not seem plausible to ask any account C to be as strong as to entail a priori that subjects that satisfy C will have genuine phenomenal states so as to be able to formulate the conceivability argument. What C should entail is that subjects that satisfy C will not be able to infer phenomenal truths from physical truths. And Chalmers' argument does not pose any problem against this possibility.

References:

- Balog, K. 1999. Conceivability, possibility, and the mind-body problem. *Philosophical Review* 108: 497-528.
- Chalmers, D. 2006. Phenomenal concepts and the explanatory gap. In T. Alter & S. Walter (eds.) *Phenomenal Concepts and Phenomenal Knowledge: Essays on Consciousness and Physicalism*. Oxford University Press. Available at: <http://consc.net/papers/pceg.html>
- Hill, C. 1997. Imaginability, conceivability, possibility, and the mind-body problem. *Philosophical Studies* 87: 61-85.
- Hill, C. & McLaughlin, B. 1999. There are fewer things in reality that are dreamt of in Chalmers' philosophy. *Philosophy and Phenomenological Research* 59.
- Loar, B. 1997. Phenomenal states. In N. Block et. al. (eds.) *The Nature of Consciousness*. MIT Press.
- Loar, B. 1999. On David Chalmers' The Conscious Mind. *Philosophy and Phenomenological Research* 59.
- Loar, B. 2003. Qualia, Properties, Modality. *Philosophical Issues* 2003: *Philosophy of Mind*.
- Papineau, D. 2002. *Thinking about Consciousness*. Oxford University Press.
- Stoljar, D. 2005. Physicalism and phenomenal concepts. *Mind and Language* 20: 469-494.
- Sturgeon, S. 2001. *Matters of Mind*. Routledge.