Physicalism and consciousness

The so-called phenomenal concept strategy (Lear, 1990) is a physicalist set of responses to anti-physicalist arguments that aims at accommodating the subjective character of consciousness to a physicalist metaphysical framework. More specifically, the strategy aims at responding to specific arguments, namely, the conceivability argument and the knowledge argument. In rough words, those well-known arguments share the same structure in that they depart from epistemic premises about the lack of deducibility of phenomenal truths from physical truths and move to the conclusion that phenomenal truths cannot be reduced to physical truths; hence physicalism must be false.

\[(\text{PHYS}) \land (P \rightarrow \neg Q)\]

Let P be the complete physical description of the world, and Q the phenomenal description of the world. The conditional is meant to express that once every minimal physical aspect of the world is settled, the phenomenal aspects of the world are necessarily settled.

The knowledge argument formulated by Frank Jackson (1982) begins by asking us to consider the lack of an a priori connection between phenomenal and physical truths. This is the ex-
planation for the reason that Mary, the brilliant scientist who knows every physical fact about visual physiology, cannot deduce phenomenal truths from physical truths. She cannot make the appropriate deductions because there is an ontological gap between the phenomenal and the physical realm. The lack of an a priori connection should entail an ontological difference between the physical and the phenomenal.

(P1) Q is not a priori deductible from P.
(P2) If Q is not a priori deductible from P, then P does not metaphysically necessitate Q.
(P3) If P does not metaphysically necessitate Q, then physicalism is false.
(C) Therefore, physicalism is false.

In the conceivability argument David Chalmers (2010) asks us to conceive of a philosophical zombie, a creature physically and functionally identical to us but with no such thing as subjective experience. To conceive of zombies (P → ¬Q) is to consider that (PHYS) is contingent, which alone violates (PHYS). Following Chalmers (2010), we can now see how the zombie argument is formulated:

(P1) P → ¬Q is conceivable.
(P2) If P → ¬Q is conceivable, P → ¬Q is metaphysically possible.
(P3) If P → ¬Q is metaphysically possible, physicalism is false.
(C) Therefore, physicalism is false.

Let P stand for all physical truths in the world and Q stand for all phenomenal truths. In (P1) the physical properties are kept constant, whereas the phenomenal properties vary. As explained previously, to conceive of a physical duplicate lacking phenomenal states is to conceive of minimal physicalism as false. Further, as conceivability implies metaphysical possibility, the metaphysical possibility of zombies is inconsistent with physicalism. The conceivability argument is clearly valid. Physicalists need to show that at least one of the premises is false. Objections to the argument will typically question the first two premises: (i) Are zombies conceivable? If they are conceivable, (ii) does it follow, as claimed in (P2), that they are possible? The proponent of the conceivability argument must answer both questions in the affirmative. Consequently, physicalists will say no to either the first or the second question. We shall now examine the possible physicalist reactions.

The phenomenal concept strategy consists in offering an alternative explanation for the lack of an a priori connection between phenomenal and physical truths that the knowledge argument and the conceivability argument hold to be true. They claim that the disconnection is not due to a property difference, but due to a conceptual difference. Mary cannot make the appropriate deductions because she lacks a phenomenal concept that can only be acquired by experience. This is because phenomenal concepts have special possession conditions. This sui generis character of phenomenal concepts is held to ensure the aposteriority of the connection between phenomenal and physical truths. If they are a posteriori connected, then the link between conceivability and possibility is broken: (P and not Q) is conceivable but needs therefore not be metaphysically possible.

So the project consists of two tasks: one to ascribe special features to phenomenal concepts, which makes them relevantly different from physical concepts, yielding a sort of conceptual dualism. The second task is to provide an alternative explanation of Mary’s ignorance. The result that type-B materialists wish to achieve is the defense of a posteriori physicalism. A posteriori physicalism holds that the relation of determination between the physical and the phenomenal is necessary but a posteriori. This can only to be done with the aid of phenomenal concepts. The idea is that phenomenal concepts have to be characterized in a way that they avoid the anti-physicalist conclusion of the conceivability argument and Kripke’s arguments against psycho-physical identification. Additionally they should provide an alternative explanation for Mary’s ignorance. Type-B materialists think that by ascribing a special feature to phenomenal concepts, a feature that makes them essentially different from physical concepts, we achieve the desiderata for responding to the anti-physicalist arguments in question.

Experience as the mark of the phenomenal

The phenomenal concept strategy aims at providing an alternative physicalist explanation of the epistemic gap posed by anti-physicalist arguments, that is, of the conceptual independence of physical and phenomenal concepts. This alternative explanation is designed so as to not expand the ontology beyond physical facts. The sui generis status of phenomenal concepts is, according to some versions of the strategy, due to their experience dependence and not due to their referring to a special realm of non-physical properties. Phenomenal concept theorists ascribe to concepts of our experience unique aspects that distinguish phenomenal concepts from other concepts, but they still refer to physical properties. One candidate criterion to separate phenomenal concepts from physical concepts is offered by the so-called experience thesis, the thesis that the possession of these concepts is experience-dependent:

Experience Thesis: One can only possess a phenomenal concept C if one undergoes the relevant corresponding experience.

The experience thesis is about concept possession. However, many versions of the strategy do not endorse the experience thesis in the sense of being a requirement for concept possession. Some phenomenal concept theorists think that experience is a requirement for special reference-fixing
mechanisms, for concept acquisition or even for involving different psychological faculties. Regardless of the specific role versions of this strategy assign to experience, all versions agree that phenomenal concepts are experience-dependent and that they are, by definition, perspectival. Unpacking the experience thesis as a guide to the conceptual independence of phenomenal concepts will help us to make the case for the phenomenal concept strategy.

Many proponents of the strategy seek to establish the existence of a binding relationship between a certain type of experience and the possession of phenomenal concepts. They claim that the only way one might possess phenomenal concepts is by undergoing an appropriate experience. In Lewis’ (2004) words: in some cases experience is the best teacher. Indeed, the phenomenal concept theorist wants to say that, in this case, experience is the only teacher: Some items of knowledge cannot be grasped unless the knower has experienced first-hand the content of concepts involved in those items of knowledge. Objective physical descriptions of these experiences are not adequate ersatz for the experience in the first-person perspective. There is something very intuitive about this. In a sense, it is a trivial claim that phenomenal concepts are tied to experience. Phenomenal concepts are concepts about experiences formed from one’s own perspective: in order for someone to attend to their own experience one needs to possess the appropriate concept, in the same way as one needs to possess appropriate concepts in order to understand thoughts, ideas etc. It is natural to strengthen the thesis that phenomenal concepts are constitutionally tied to experiences, that is, experience is a necessary condition for one to come in possession of phenomenal concepts. If phenomenal concepts are special because they have special possession conditions, then the strategy might mobilize this characteristic of phenomenal concepts to neutralize the epistemic arguments. Stoljar (2005) adduces two independent arguments against the phenomenal concept strategy’s response to the conceivability argument on the one hand and to the knowledge argument on the other. First, against the response to the conceivability argument, Stoljar argues that no special feature of phenomenal concepts can ensure the a posteriority of physicalism. Second, against the response to the knowledge argument, he argues that the thesis that Mary cannot make the appropriate deductions because she lacks the relevant concepts is false. In the next two sections I will consider these objections in turn. Let us consider the psychophysical conditional (PHYS) again:

\[(PHYS) \Box (P \rightarrow \neg Q)\]

Stoljar characterizes the central project of the phenomenal concept strategy as that of explaining how the a posteriority of the conditional is possible without running into Kripkean difficulties. So, in order for the phenomenal concept strategy to succeed, the experience thesis or any replacement thesis to the same effect must entail the a posteriority of (PHYS). Nevertheless, he will argue that the experience thesis does not entail the a posteriority of the psychophysical conditional, hence the strategy will fail.

As discussed previously, there are facts about the nature of phenomenal concepts which make them different from physical concepts. Each specific account of phenomenal concepts assigns different properties to phenomenal concepts. Although not all versions of the phenomenal concept strategy endorse the experience thesis as formulated in the previous chapter, Stoljar’s objection is not restricted to the experience thesis, rather it is aimed at all versions of the phenomenal concept strategy.

The a priori physicalism and phenomenal concepts: The a priori synthesizable objection

Stoljar’s first objection to be considered applies specifically to the treatment of the conceivability argument by the phenomenal concept strategy. He begins by distinguishing between two ways of understanding the conditional: the a priori and the a priori synthesizable (Stoljar, 2005).

The a priori:

\[A \rightarrow B\] is a priori if a sufficiently logically acute person who possesses only the concepts required to understand it is in a position to know that it is true (478).

The a priori synthesizable:

\[A \rightarrow B\] is a priori synthesizable if a sufficiently logically acute person who possessed only the concepts required to understand its antecedent is in a position to know that it is true (478).

To illustrate this distinction, consider (1):

\[(1)\] If \(y\) is rectangular, then \(x\) has some property or other (478),

where (1) is clearly a priori, but not a priori synthesizable, since a logically acute person who knows only the concepts involved in the antecedent of the conditional—the concept Rectangular—is not in a position to a priori synthesize the consequent. A logically acute person who lacks the concept Property cannot understand the consequent of the condition; hence, she is not in a position to understand the conditional. Being a priori synthesizable entails being a priori, but not the other way around; a priori conditionals may fail to be a priori synthesizable as in (1).

To avoid the conceivability argument, the phenomenal concept strategist must hold that the conditional (PHYS) is not a priori. The crucial premise of the phenomenal concept strategy is that the experience thesis entails that (PHYS) is not a priori. This is because, for physicalism to be true, P
must entail Q. Since we concede the epistemic gap of the anti-physicalist arguments, the entailment is not a priori, and the only option left is to show that the conditional is a posteriori. The experience thesis is crucial because it marks the cognitive difference between phenomenal concepts and physical concepts. So, the project of the phenomenal concept strategy is to argue that because phenomenal concepts are experience-dependent, they are conceptually independent from physical concepts; hence the conditional is a posteriori. Being a posteriori, however, is equivalent to being not a priori, but not equivalent to being not a priori synthesizable. The phenomenal concept strategy claims that the experience thesis entails that (PHYS) is not a priori. Stoljar, however, wants to show that the experience thesis actually entails that (PHYS) is not a priori synthesizable. Not being a priori synthesizable is irrelevant for the a posteriori of (PHYS). It is irrelevant because one can find many examples of a priori propositions like (1) which are not a priori synthesizable but still a priori. One can say that all that the experience thesis entails is that (PHYS) is not a priori synthesizable. Granting that (PHYS) is not a priori synthesizable does not eliminate the possibility that (PHYS) is not a priori, like in (1). So there seems to be a logical gap in the suggestion that the experience thesis tells us that the conditional is a posteriori. What we wanted was a reason to suppose that it was not a priori. What we have is a reason to suppose that it is not a priori synthesizable (479). When failing to point out that (PHYS) is not a priori, the experience thesis does not provide an answer to the conceivability argument.

Proponents of the strategy could object to this line of reasoning by saying that examples like (1) are relevantly different from (PHYS), because (1) does not involve phenomenal concepts, while (PHYS) connects the antecedent which contains only ordinary concepts and the consequent which contains phenomenal concepts. Thus, the distinction between a priori and a priori synthesizable could be ignored. But Stoljar has a follow-up argument to the same conclusion. Stoljar asks us to consider (2), a statement which is relevantly like the conditional in that it connects an antecedent containing an ordinary concept with a consequent containing a phenomenal concept (possession of the concept that appears in the consequent of the conditional requires experience):

(2) If x is a number, then x is not a red sensation (479).

Sentence (2) is, like (1), clearly a priori, but not a priori synthesizable. Someone who lacks the concept red sensation required to understand the consequent of the conditional cannot deduce the consequent from the antecedent. The problem for the physicalist is if (PHYS) turns out to be a case like (2): clearly a priori, but not a priori synthesizable. Then, the physicalist would have failed to show the a posteriority of (PHYS) through the experience thesis. The lack of a priori synthesizability would explain the conceptual independence between phenomenal concepts and physical concepts. But, in this case, mere conceptual independence does not lead to lack of a priori connections.

The phenomenal concept theorist could consider that the distinction works in her favor, at least in the case of the knowledge argument: Mary’s inability to deduce Q from P is not explained by claiming that Q and P refer to different properties, but it is explained by the fact that Mary simply cannot synthesize those truths a priori. She cannot a priori synthesize Q from P because she lacks some crucial concepts to understand the consequent of the conditional (a phenomenal concept). However, Stoljar advances an independent objection against the strategy’s treatment of the knowledge argument (which will be addressed at the end of this section). For now, the distinction above is aimed at the strategy’s response to the conceivability argument.

The proponent of the phenomenal concept strategy could argue that Stoljar’s objection works only against versions of the strategy that are committed to the experience thesis. Do other versions also fail in view of Stoljar’s distinction between the a priori and the a priori synthesizable? Stoljar analyzes the version proposed by Hill and McLaughlin (1999) to show that not only the experience thesis, but any replacement thesis fails in view of the distinction.

According to Hill and McLaughlin (1999), phenomenal concepts and physical concepts are governed by different epistemic constraints and presuppose use of different faculties. One difference between phenomenal concepts and physical concepts is that the former are self-presented:

Self-presentation thesis: It is a conceptual truth that if I have a red sensation, and if I have the concepts and focus my attention on the matter, I will thereby come to know that I am having (Stoljar, 2005, p. 483).

On the other hand, “it is not a conceptual truth that if I am in some overall physical conditional P, and if I have the concepts and focus my attention on the matter, I will thereby come to know that I am in P” (Stoljar, 2005, p. 483) This is to say that, if there is a conditional whose antecedent contains an ordinary concept and the consequent contains a self-presenting concept, then the conditional cannot be a priori. Stoljar grants that phenomenal concepts might be self-presenting in the above sense. Still, he holds that the self-presentation thesis, like the experience thesis, cannot explain the posteriority of the psychophysical conditional. A central reason to think that (2) poses a problem for the phenomenal concept theorist is to assume that applying the negation of a phenomenal concept not a red sensation requires possession of the phenomenal concept red sensation; hence the possession of negations of phenomenal concepts are experience-dependent, or, according to Hill and McLaughlin’s version, the negation of concepts of experience requires the possession of self-presenting concepts. So not a red sensation is self-presenting just as red sensation is. On the other hand, possession of the concept number is clearly not self-presenting (nor experience-depen-
The sentence (2) contains theoretical concepts in the antecedent and self-presenting concepts in the consequent just like in (PHYS). If the reason that (PHYS) is a posteriori and appears to be contingent in Hill and McLaughlin’s version of the strategy is that these two different kinds of concepts are entailed, then (2) should also be a posteriori and have an appearance of contingency. Nevertheless, (2) is clearly a priori, so neither the self-presenting thesis nor the experience thesis can be correct.

The physicalist can respond to Stoljar’s observations by claiming that cases like (2) (If x is a number, then x is not a red sensation [479]) are clearly a priori because the negation of a phenomenal concept such as not a red sensation does not require possession of the phenomenal concept red sensation. Mary, who lacks all kinds of color-related phenomenal concepts, including red sensation, is in a position to know a priori that (2) is true if she possesses at least partial understanding of the consequent, that is, if she possesses the concept sensation. She knows that, if something is a number, then it is definitely not a sensation of any kind. She may also possess those concepts second-handedly (through testimony). She may know that sensations are usually associated with perceptual states, while numbers are not, and that should be enough for Mary to know (2) a priori. Stoljar’s reasoning does not pay due attention to the disanalogy between (PHYS) and (2). There are many ways to possess the negation of a phenomenal concept which does not involve itself the acquisition of a phenomenal concept. It is perfectly plausible to possess a nonphenomenal concept about a phenomenal concept (without possessing the phenomenal concept in question). I can refer to other people’s sensation without sharing their sensations. The concepts that are formed through testimony are not phenomenal concepts.

The disanalogy between cases like (2) and the psychophysical conditional may be brought out by considering analogous strategies concerning other kinds of concepts. Consider, for example, a natural kind concept strategy applied to the following conditional involving a natural kind term (see Parvin, 2008).

(3) If x is Au then x is gold.

The sentence (3) is necessary and a posteriori because it connects a theoretical concept (Au) to a natural kind concept (gold). But now consider:

(4) If x is a number then x is not gold.

(4) connects a theoretical concept to the negation of a natural kind concept, but (4) is clearly a priori. Although the negation of a natural kind is not a natural kind, the question is whether the negation of a natural kind concept requires possession of a natural kind concept. According to Stoljar’s thesis, the natural kind concept gold is required to understand the consequent of (3) but not the antecedent, just as it is with (4). But (4) is a priori like (2), and thus, if we were to follow Stoljar’s reasoning, we would conclude that this treatment of natural kind terms fails.

Another disanalogy between cases like (PHYS) and other conditionals which connect consequents containing experience-dependent concepts and antecedents containing only ordinary concepts is (5) (Diaz-Leon, 2008):

(5) If x is a square circle then x is a red sensation.

Because the antecedent is a priori false, the whole sentence is a priori false. There is no need to understand the consequent in this case. This illustrates another type of sentence structurally like (PHYS) but with a different epistemic status. The fact that (4) and (5) are a priori does not undermine the a posteriority of (PHYS), it shows only that we do not use the same criteria to evaluate the epistemic status of (PHYS) that we use to evaluate (4) or (5). And this is showed by Diaz-Leon (2008):

All that the phenomenal concept strategy needs is that phenomenal concepts are not a priori connected with the physical-theoretical concepts that appear in the antecedent of the psychophysical conditional [...]. The strategy does not need to show that no phenomenal concept is a priori connected to any physical-theoretical concept. If there are plausible reasons for thinking that some phenomenal concept is a priori connected with some physical-theoretical concept, why should the advocate of the phenomenal concept strategy not accept that? (Diaz-Leon, 2008, p. 608).

It seems that the strategy achieves the goal of ensuring the aposteriority of the conditional by postulating the dependence of phenomenal concepts vis-à-vis experiences. This dependence grants the aposteriority of (PHYS) without running into Kripkean problems for psycho-physical identification (or, in our cases, entailment).

**Experienced Mary**

The second part of Stoljar’s argument concerns the treatment of the knowledge argument by the proponents of the strategy. Stoljar thinks that the knowledge argument can be reformulated such that the new concept explanation is not a satisfactory answer to the problem. Versions of the phenomenal concept strategy that are committed to the experience thesis explain Mary’s epistemic progress in terms of the acquisition of a new concept. Only after leaving the room can Mary acquire the new concept that enables her to make the appropriate deductions. At this point, the distinction introduced by Stoljar to work against the strategy’s response to the conceivable argument could work in favor of the strategy’s treatment of the knowledge argument. The experience thesis
explains why phenomenal truths are not a priori synthesizable from physical truths, and that would be enough to explain the epistemic gap between phenomenal concepts and physical concepts. However, Stoljar rejects this line of reasoning arguing that there is an independent reason to reject the phenomenal concept theorist’s explanation of Mary’s ignorance. Stoljar offers a different thought experiment in order to show that even if Mary possessed the relevant phenomenal concepts, she would still not be able to deduce phenomenal truths from physical truths. Since the phenomenal concept strategy’s treatment of the knowledge argument is to explain Mary’s ignorance in terms of her lacking the phenomenal concept, the strategy would fail.

Stoljar asks us to consider experienced Mary. She is just like Mary for the first part of the story; after experienced Mary is released from the black and white room, she has color experiences and, because of that, she is able to apply the relevant phenomenal concepts. She is, later, recaptured and returned to her room. After she returns to her room, experienced Mary undergoes a process of selective amnesia: she forgets the correct application of phenomenal concepts. She still knows what it is like to see green, thus she still possesses the phenomenal concepts that she acquired during her short period of freedom. However, she fails to make associations like “looking at Granny Smith apples typically causes green sensations” or “having arthritis causes pain”. Experienced Mary cannot deduce phenomenal truths from physical truths even though she possesses the corresponding phenomenal concepts. In Stoljar’s view, she knows the antecedent of the conditional (PHYS) but not its consequent. Stoljar wants to show that this new scenario turns the acquisition of phenomenal concepts irrelevant to explain Mary’s ignorance.

I think that the tale of experienced Mary undermines the fundamental premise of the knowledge argument against physicalism, viz. that Mary has complete physical knowledge of visual physiology. If a priori physicalism were correct from Mary’s complete physical description of visual physiology, Mary would be able to deduce phenomenal truths, for example, ‘what it is like to see red’. But this leads us to the second premise of the argument; Mary cannot deduce phenomenal truths from physical truths. This epistemic gap puts physicalism under pressure. The physicalist needs to explain how to accommodate this lack of a priori deducibility in the thesis that physical truths metaphysically necessitate phenomenal truths. In Stoljar’s second scenario, Mary leaves the room for the first time and gains knowledge to understand all the conceptual framework necessary to understand the physical knowledge she has while inside the room, but then she is recaptured and forgets the correct application of the concepts she had just learn how to apply. What seems to be missing in this version of Mary is information which belongs to the antecedent of the conditional P→Q, not to the consequent. Deleting part of her memory which would enable Mary to correctly apply her color concepts is to delete information which belongs to the physical knowledge. Thus, it turns the physical knowledge of experienced Mary incomplete; so, experienced Mary would not possess the relevant concepts to understand the antecedent of the conditional. Without means to understand the antecedent, Mary cannot understand the conditional. Stoljar would only be able to conclude that the acquisition of phenomenal concepts is irrelevant to explain Mary’s ignorance if she were allowed to keep her memories as she returned to the black and white room.

The correct application of phenomenal concepts would be available to experienced Mary, since it is, in a sense, information belonging to the objective, physical domain. Mary must not undergo any experience to know that a deep cut in one’s skin typically causes pain or that looking at Granny Smith apples typically causes green sensation, and that looking at a red fire hydrant typically causes red sensations. Original Mary possesses all this knowledge inside the room, which still does not enable her to deduce phenomenal truths from physical truths. If we add new phenomenal concepts to Mary’s set of beliefs, but later discard her beliefs which rule the application of those concepts, we discard knowledge that is crucial for Mary to understand the antecedent of the conditional but not the consequent, this is knowledge that used to be part of her complete physical description about the world. Experienced Mary gains phenomenal concepts, but loses ordinary concepts. Experience would only enable Mary to master the concepts she possesses partially while still inside the room. However, if experienced Mary already possessed such concepts, the case Stoljar presents to us is not a case in which she knows the antecedent of the conditional but cannot deduce the consequent. It is a case in which information belonging to the antecedent is omitted. Mary cannot deduce phenomenal knowledge from incomplete physical knowledge. At this point one could ask why would the examples of correct phenomenal belief application be considered physical knowledge given that pain is a phenomenal concept. It is simply by distinguishing between phenomenal concepts, nonphenomenal concepts about phenomenal concepts and psychological concepts that we can understand that this is not at all controversial. The phenomenal concept red sensation is the concept of the specific type of sensation someone typically has when looking at red things. It is very different from the concept Red, for this concept refers to red things. It is also different from the concept The sensation caused by red things as certain color-blinds will feel green upon looking at red things (Stoljar, 2005). It is also important to differentiate phenomenal concepts from psychological concepts (Balog, 2009). The latter are analyzed in functional terms. A psychological concept is an objective, third-person concept, since it refers to a mental state, which contains no reference to the subject’s subjective experience, whilst a phenomenal concept is formed from one’s own perspective. This is just to say that not all concepts about phenomenal states are phenomenal concepts. One can have nonphenomenal concepts about phenomenal states, just as color blinded people believe that Granny Smith apples are green without ever having experienced colors.
Conclusion

In order to respond to the conceivability argument, the proponent of the strategy had to argue for the a posteriority of (PHYS) by appealing to the experience thesis or any replacement thesis drawing on the perspectival nature of phenomenal concepts. The distinction between a priori statements and a priori synthesizable statements is introduced by Stoljar in order to show that the best that the experience thesis accomplishes is to show that physicalism is not a priori synthesizable. The latter is admittedly irrelevant to the explanation of why (PHYS) is conceivable as false, but metaphysically impossible. Nevertheless, Stoljar ignores a clear disanalogy present between conditionals like (PHYS) and other cases involving ordinary concepts. If Stoljar’s argument worked against the phenomenal concept strategy, it would also generate undesirable results for the way we think about natural kind concepts and for paradigmatic cases of a priori conditionals. Besides, I also dispute the assumption made by Stoljar that negation of a phenomenal concept require possession of that concept.

Furthermore, I intended to show that the tale of experienced Mary is a non-starter. His strategy consists in deleting part of Mary’s memory to show that even if she possesses the relevant phenomenal concepts but loses the ability to correctly employ them, she is not able to deduce phenomenal truths from physical truths. Deleting part of her memory which would enable Mary to correctly apply her color concepts is to delete information which belongs to the antecedent of the conditional. Without knowledge of the antecedent, she cannot understand the consequent.

References


Submitted on November 14, 2017
Accepted on December 13, 2017