

Leaving the Dream Behind: Why the Metaphysics of Consciousness cannot be Unveiled by Conceivability Arguments

Luis Alejandro Murillo-Lara

Augustinian University, Universidad Nacional de Colombia

Email luisalejandro.murillo@gmail.com

ORCID <https://orcid.org/0000-0002-8156-4128>

Abstract. In this paper, I present a novel objection to Chalmers’s “master argument” against the privileged strategy of ‘type B’ physicalists to account for the explanatory gap (the “phenomenal concepts strategy”). Specifically, I argue that the second horn of the master argument gets wrong why zombies cannot have our epistemic situation with regard to consciousness. Zombies cannot have a kind of mental state that we have. If something must have all of our psychological attributes to share our epistemic situation, then zombies cannot serve the purpose of the second horn of the dilemma. By way of background, I begin by briefly outlining a related argument against physicalism, also advanced by D. Chalmers – the “conceivability argument.” I highlight some of the primary challenges with this argument and present additional criticisms. Finally, through a brief examination of panprotopsychism, I consider what lies ahead if Chalmers’s arguments are conceded. I conclude that the phenomenal concept strategy is a sound explanation for why the conceivability of zombies likely does not imply their metaphysical possibility.

Keywords: master argument, phenomenal concepts strategy, conceivability, possibility

Atsisveikinant su svajone: kodėl suvokiamumo argumentai nepajėgia atskleisti sąmonės metafizikos

Santrauka. Šiame straipsnyje pateikiu naują prieštaravimą Chalmerso kertiniam argumentui prieš „B tipo“ fizikalistų privilegijuotą strategiją pagrindžiant aiškinimo spragą („fenomenalių sąvokų strategija“). Aš konkrečiai teigiu, kad antrasis kertinio argumento dėmuo neteisingai interpretuoja, kodėl zombiai negali patekti į mūsų episteminę situaciją sąmonės atžvilgiu. Zombiams negali būti būdingos tokios mentalinės būklės, kurias turime mes. Jeigu visi mūsų psichologiniai atributai yra būtina mūsų episteminės situacijos sąlyga, tai zombiai negali pasitarnauti kaip antrasis dilemos dėmuo. Kalbėdamas apie kontekstą, aš visų pirma trumpai aptariu susijusį argumentą, nukreiptą prieš fizikalizmą, kurį taip pat naudojo D. Chalmersas – „suvokiamumo argumentą“. Aš akcentuoju kai kuriuos svarbiausius iššūkius šiam argumentui, kartu pateikdamas ir papildomų kritikos elementų. Galiausiai, glaudžiai panagrinėjęs panprotopsichizmą, aš svarstau, kas mūsų lauktų, jei priimtume Chalmerso argumentus. Prieinu išvadą, kad fenomenaliųjų konceptų strategija yra patikimas paaiškinimas, kodėl zombių suvokiamumas veikiausiai nereiškia jų metafizinės galimybės.

Pagrindiniai žodžiai: kertinis argumentas, fenomenalių sąvokų argumentas, suvokiamumas, galimybė

Received: 19/06/2023. Accepted: 02/10/2023

Copyright © Luis Alejandro Murillo-Lara, 2023. Published by Vilnius University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

There are several arguments against physicalism: zombies (Chalmers 1996), cloistered scientists (Jackson 1982), and lots of qualia types (absent, inverted, and even dancing and fading). What they have in common is that they present a conceivable situation and thence conclude that physicalism is false, either because physicalism predicts that this situation is impossible or because physicalism is insufficient to account for such a situation. They can thus be called “conceivability arguments” or “the conceivability argument,” singular, to refer to the common formal structure. The conceivability argument has been refined over the years and has gained its most mature form in the work of David Chalmers (2002, 2010), according to whom the conceivability-possibility connection is not unrestricted.

The response of most physicalists to conceivability arguments has been consistently the same: if those scenarios are conceivable, their conceivability does not entail that they are metaphysically possible. A popular view among those theorists is that, in fact, those scenarios are conceivable – that is, that there is an explanatory gap between P and Q above, allowing to conceive ($P \ \& \ \neg Q$). The explanatory gap would arise because we use two independent sets of concepts (P and Q) to grasp the same reality (so that the gap is epistemic and without metaphysical consequences) (e.g., Hill 1997; Loar 1997; Papineau 2006). Chalmers (2003a) labels as ‘type B physicalism’ the acknowledgment of the explanatory gap while endorsing physicalism, and the aforementioned account of the explanatory gap ‘the phenomenal concept strategy’ (PCS) (Chalmers 2006).

I will begin with a few words about Chalmers’s version of the conceivability argument, since it is the conceptual background for his rejection of the PCS. The concept of a philosophical zombie and the claim that zombies are at least ‘negatively’ conceivable first appear in his argument that conceivability entails possibility and are key elements of Chalmers’s argument against the PCS. The intention of that section is, of course, far from exhaustive. Given my primary goal and space limitations, I will only lay out the basics of the argument, as well as some serious difficulties it faces.

The core of the manuscript is devoted to developing a novel objection to Chalmers’s (2006, 2010) critique of PCS. That critique takes the form of a dilemma, the second horn of which I will focus on. This horn depends on the claim that zombies do not have our epistemic situation with regard to consciousness. However, I will argue, that is the case because of a kind of the mental state that they cannot have. In turn, the physicalist thesis can be expressed as follows: to have our epistemic situation with regard to consciousness, something must be physical and have all the mental states we have. If we add to this the doubts raised about the first horn, the two horns can be defused, and the dilemma disappears.

Finally, in order to highlight the dead ends to which Chalmers’s broad perspective leads to, I will briefly engage with a purported contender for the theory of consciousness privileged by Chalmers (1996, 2010, 2015, 2017). My conclusion will be that the time for overcoming conceivability arguments and their kin is overdue.

The Conceivability Argument: A Frail Titan

Let us begin by introducing Chalmers's (1995, 2002, 2010) version of the conceivability argument in a nutshell: a complete microphysical description of the actual world (P) is compatible with the absence of any phenomenal truth (Q); in other words, $(P \ \& \ \neg Q)$ is *a priori* coherent. From that, it follows that there are epistemically possible scenarios in which $(P \ \& \ \neg Q)$ is true – scenarios where the sentence's 'primary intension' is true. In addition, Chalmers maintains that, for every epistemically possible scenario that is 'ideally', 'positively', and 'primarily' conceivable, there is a metaphysically possible world (PW) in which the corresponding sentence is true. In particular, it is the 'semantic stability' of P and Q what makes their 'primary' and 'secondary' intensions to be the same, whereby the primary and secondary intensions of $(P \ \& \ \neg Q)$ will be identical as well – otherwise put, since it is epistemically or 'primarily' possible that the sentence at hand is true, it is also 'secondarily' possible. If so, physicalism is not only explanatorily insufficient but metaphysically false.

Regarding the connection between conceivability and possibility, Chalmers distinguishes a number of forms of conceivability: for starters, *prima facie* conceivability from that produced by undefeatable reasoning ('ideal'). A sentence is 'negatively' conceivable if it does not involve contradiction (if its negation is not *a priori*), whereas a product of 'coherent modal imagination' makes a sentence 'positively' conceivable. Finally, a sentence is primarily conceivable when its primary intension is true (that is, when it is conceivable for everything that we know *a priori*) and secondarily conceivable when its secondary intension is (that is, when it is conceivable given our *a posteriori* knowledge). In Chalmers's view, $(P \ \& \ \neg Q)$ is not only negatively conceivable (as demonstrated by the existence of an explanatory gap between P and Q), but ideally so, and its primary intension is true. Chalmers (2002) also claims that whatever is ideally negatively conceivable is ideally positively conceivable as well (2002: 158). Therefore, $(P \ \& \ \neg Q)$ is metaphysically possible.

Regarding 'semantic stability', Chalmers considers a materialist reply, according to which, *secondary* (i.e., not primary) conceivability is a guide to metaphysical possibility. Chalmers's answer is that the primary and secondary intensions of P might be identical, as happens with Q (2002: 185-186). Think of some microphysical description of a phenomenon: it seems likely that the same underlying item is the one responsible for the phenomenon in another PW – if so, its *epistemic* and *counterfactual* intensions would be the same ('semantic stability'). If P and Q are semantically stable and $(P \ \& \ \neg Q)$ is primarily conceivable, then its secondary intension is also true.

There are several drawbacks preventing Chalmers's version of the conceivability argument against physicalism from working (see Hill and McLaughlin 1999; Loar 1999; Yablo 1999; Schroeter 2004). Before moving on to the core of the manuscript, I want to briefly highlight three of those drawbacks, insofar as they attack the key elements for the passage from conceivability to metaphysical possibility (the notions of primary, ideal, and positive conceivability). The third, and part of the second, are my own contributions. If

those concerns are on the right track, even in Chalmers's sophisticated version, conceivability arguments are unable to draw metaphysical conclusions:

1. 'Primary intensions' probably cannot be considered a kind of meaning

Block and Stalnaker (1998) note that, given an expression and a set of PWs in which it has semantic value, a function will take those PWs into the semantic values. If semantic values are 'secondary intensions', such a function would be a 2D intension determining a 'primary intension' (Block and Stalnaker 1998: 38-39). So, even if we do not know the semantic value of an expression, we know its primary intension – there can always be a function that takes a world into the semantic value an expression has in that world. Thus, they doubt that Chalmers's 2D intensions have to be considered a kind of meaning, rather than a representation of possible meanings. Chalmers's probable answer is that the primary intensions of concepts are not determined this way and that it is rather the concept 'as we use it' that determines its meaning in other PWs (see Chalmers 1996: 366). This answer is unpromising, because the meaning we have in mind when considering another PW cannot be the secondary but the primary meaning, which should then already be determined.

Block and Stalnaker (1998) underscore that our thinking about the adequate epistemic responses in counterfactual situations is not only shaped by our concepts. It is also influenced by theories and methodological principles that are *a posteriori*. Besides, Chalmers's primary intensions are initially assigned on the basis of empirical experience, that is, on patterns of understanding that are *a posteriori*. If so, there might be no way of entertaining a scenario as actual that does not rely on any empirical assumptions.

2. Problems with ideal reasoners

Another obstacle for Chalmers's view comes from the status of ideal reasoners. Fraga (2017) argues that if ideal reasoners have cognitive limitations, they cannot (deductively) rule out *a priori* $(P \ \& \ \neg Q)$. Alternatively, in order to (deductively) rule out *a priori* that $(P \ \& \ \neg Q)$, ideal reasoners should have no cognitive limitations. Now, that a 'finite' ideal reasoner does not (nondeductively) rule out *a priori* $(P \ \& \ \neg Q)$ is a nonconclusive reason for $\diamond(P \ \& \ \neg Q)$ (2017: 51). On the other hand, Fraga emphasizes that if ideal reasoners do not have cognitive limitations, they would not be of use to guide us about the rational patterns of inference we, finite reasoners, are supposed to accept.

This point can be pressed further. Reasoners without cognitive limitations not only would be useless to know what reasoning patterns we should accept: it is contentious that we (finite reasoners) could even ascertain what is conceivable for them. Not only could their memory potentially store infinite information, but they need not have the same set of cognitive skills and patterns of reasoning that we have. If so, that sentences as $(P \ \& \ \neg Q)$ are conceivable for reasoners without cognitive limitations is simply unwarranted – while, as Fraga notes, being conceivable for reasoners with cognitive limitations is inconclusive for $\diamond(P \ \& \ \neg Q)$.

3. Positive conceivability as a better guide to possibility

Positive conceivability, which Chalmers defines as coherent modal imagination, is said to go beyond a pictorial representation of something (2002: 149-150). It allows one to positively conceive some scenarios that one could not directly perceive (say, the nucleus of an atom splitting) because it is imagination ‘accompanied by interpretation and reasoning’, so that we manage to conceive scenarios by picturing something *we take to* verify some statement (2002: 149, 150, 152, 164). Positive conceivability thus seems to allow imagining the hardest things: imagining a pellet breaking up could count as imagining the nucleus of an atom splitting as long as the right interpretation accompanies the image or as long as one takes that image to verify the relevant statement. Nevertheless, all that is perplexing: Are there any limits to the interpretation of an image, i.e., to this taking an image to verify something? If there are not, the image plays no substantial role – the whole burden is on the interpretation. If what matters for something to be positively conceivable is an attitude towards a statement, for it to be *coherent* it is not enough that it is not contradictory (that is enough for *negative* conceivability). In this sense, we seem to lack an adequate, sound picture of positive conceivability.

Chalmers’s version of the conceivability argument holds that whatever is ideally, primarily, and positively conceivable is metaphysically possible. There are serious reasons to doubt that claim: the notions of primary conceivability and positive conceivability on which the claim is built are obscure, whereas ideal conceivability does not allow to draw the desired conclusion. On the one hand, it is unclear how ‘primary intensions’ are determined or whether they are truly *a priori*. That also thwarts the version of the conceivability argument particularized for $(P \ \& \ \neg Q)$, the argument from semantic stability. On the other hand, we might not be entitled to claim that $(P \ \& \ \neg Q)$ is ideally conceivable if ideal reasoners have no cognitive limitations; if they do have those limitations, conceiving this sentence cannot be a conclusive reason for the possibility of $(P \ \& \ \neg Q)$. Finally, positive conceivability contains an extremely liberal notion of imagination, whereby a more rigorous definition is needed. (To the best of my knowledge, answers to these issues are nowhere to be found in Chalmers’s work). As I said, Chalmers’s conceivability argument serves as the background against which his attack on the PCS is built. As we will see, the notion of a philosophical zombie (i.e., a being for whom $[P \ \& \ \neg Q]$ is true) is crucial to such an attack, as is the claim that P does not entail Q (i.e., that $[P \ \& \ \neg Q]$ is negatively conceivable). In the remainder of the paper, I shall argue that Chalmers’s attack on the PCS suffers a similar fate to his conceivability argument.

Phenomenal Concepts And the Master Argument

As noted at the beginning, some defenders of physicalism concede that $(P \ \& \ \neg Q)$ is conceivable (‘type B physicalism’) but argue that its conceivability does not entail that it is metaphysically possible. There is an explanatory gap between P and Q without

metaphysical consequences because we use two independent sets of concepts (P and Q) to grasp the same reality ('the phenomenal concept strategy' [PCS]). Diversity among proponents of this account of the explanatory gap comes, among others, from specificities in the way they characterize 'phenomenal concepts' and the way they differ from 'physical concepts' (see Hill 1997; Loar 1997; Papineau 2006). Unsurprisingly, Chalmers (2006) has objected against every version of the PCS and has made it at once. According to what Chalmers calls 'the master argument', any version of the PCS faces a dilemma: either the right account of phenomenal concepts (C) cannot be explained in physical terms (P), or it cannot explain our epistemic situation with regard to consciousness (E). The first horn follows from granting that P does not entail C – we would have an explanatory gap between them. The second horn comes from claiming that P does indeed entail C – that would mean that zombies satisfy C despite (by hypothesis) not sharing E with us, making C inadequate to account for E .¹ Summarily, any version of the PCS will be either explanatorily inadequate or inexplicable in physical terms.

In this section, I want to dwell on the master argument by focusing on the second horn and then commenting on the first horn. A widely accepted strategy for addressing the second horn is to claim that zombies might be related not to phenomenal properties, but to some kind of homologous properties (that are not conscious, but 'schmonscious' [Papineau 2006: 140-143]). If so, there would be no substantial asymmetry between the relationship (a) phenomenal concepts-phenomenal states (in our case) and (b) 'phenomenal' concepts-schmonscious states (in the zombie). Thereby, P might entail C , and C might entail E because the latter need not include only phenomenal concepts (it might include those that refer to schmonscious properties). Nevertheless, a proponent of the master argument may object that the strategy violates the zombie hypothesis (i.e., that, *ex-hypothesi*, zombies are beings with nothing epistemically analogous to consciousness [Chalmers 2006: 186]). Thus, since the proponent of the master argument is not forced to concede them, the proponent of schmonsciousness must first motivate the existence of the related properties on an independent basis. Most importantly, such a strategy runs the risk of leaving the master argument untouched: Chalmers's point is that zombies lack our inner lives, but schmonscious properties say nothing substantial about how conscious properties relate to E (or to physical/functional properties) beyond the claim that the relationship is *mirrored* by another (Papineau 2006: 141; Chalmers 2006: 187).

I agree that the second horn of the master argument does not work, but for a different reason. So, in what follows, my goal is to provide a novel objection to the second horn (and perhaps a novel take on the dilemma more generally). Both horns would be equally flawed: the conception of phenomenal beliefs that is presupposed in anti-physicalist arguments blocks the second horn (opening a way out of the dilemma for the physicalist). Meanwhile, based on a notion of explanation the PCS does not need to accept, the first horn demands from the PCS something it admits cannot be done.

¹ In Chalmers's (2006: 177) view, E comprises the truth-values and epistemic status of an individual's beliefs. Along these lines, a zombie's phenomenal beliefs differ from ours in their truth-value or epistemic status.

Chalmers's claim, made at the opening of the second horn, that the phenomenal beliefs of zombies differ from ours in their truth-value (2002: 177) needs refining. Phenomenal beliefs are about one's phenomenal states, like the belief that one is in pain (Chalmers 2003b, 2010). When one is not in pain, one's belief "I'm not in pain" is true – as is the zombie's belief that it is not in pain. Accordingly, zombies might have a lot of true phenomenal beliefs, namely, those that deny the self-attribution of phenomenal states. Perhaps not all of the zombie's beliefs self-attributing phenomenal states would differ from ours in their truth-value – if one is not in pain, both its belief and one's "I'm in pain" will be false. Thus, only our beliefs truly self-attributing phenomenal states differ in their truth-value from the zombie's beliefs self-attributing phenomenal states.²

What prevents zombies from sharing *E* with us is therefore a subset of the beliefs called 'phenomenal'. What makes a belief belong to that subset is the existence of a phenomenal state justifying it – whose absence makes false the zombie's beliefs self-attributing phenomenal states. The kind of justification in play here does not seem to be inferential, but the one sometimes attributed to perceptual beliefs – those are justified by perceptions, which, in turn, are caused by objects (Silins 2021). Phenomenal beliefs would be analogously justified by phenomenal states. If so, (true) positive phenomenal beliefs are essentially tied to their causal history – at least to a causal link, namely that of the phenomenal state and its cause. Nevertheless, that would mean that zombies cannot have those beliefs, since objects do not cause them to have phenomenal states (which, in turn, prevents the beliefs at stake from arising).

Someone might say that we have lost our way at some point: nothing seems to prevent the zombie to have the belief "I'm in pain," perhaps with the same inferential connections of one's belief that one is in pain. Let us set aside for the moment the issue of how two beliefs with different truth-values can have the same inferential profile. What matters is that this is not the view on phenomenal beliefs anti-physicalist arguments use: the reason why Mary, before leaving her black and white room, cannot have true, positive phenomenal beliefs about her experiences of red is that she has not had any of those experiences – and she has not had them because she has not been in contact with any red surfaces (Jackson 1982). If, regardless of its truth-value, its role in the belief network were what made a zombie's belief to be phenomenal, then there would be no difference between the zombies' phenomenal beliefs and mine (preventing them to satisfy *E*). If, as Chalmers (2006: 183) emphasizes, in order to satisfy *E* truth-value matters, it seems that causal relationships matter as well – and then maybe zombies strictly cannot have some phenomenal beliefs, since they cannot instantiate the relevant causal relationships essential to them.

² The case might be challenged where both our phenomenal beliefs and those of the zombies are false, because it might require that someone sincerely believes that they are in pain despite not actually being in pain. It seems difficult to maintain such a challenge independently from an infallibilist, privileged-access view on our phenomenal beliefs – it amounts to claiming that whenever a person believes "I'm in pain," they are right. At any rate, it would remain that only the zombie's beliefs self-attributing phenomenal states would differ from ours in their truth-value. In this sense, it is worth noting that, later on, Chalmers himself restrains his claims to "positive truths" (e.g., 2015: 249).

Let us imagine a physicalist who maintains that P indeed entails C . Remember that physicalism can be described as the thesis that everything supervenes on the physical – including the psychological (Ψ) (see Kim 1993). This is not disputed by Chalmers, whose quarrel is rather with the supervenience of the phenomenal (Q) on the physical (see Chalmers 1996).³ That physicalist's view can be described as the claim that, in order to satisfy C , something has to satisfy P and Ψ (let us call such a being our 'true doppelganger'). The above discussion points to the claim that zombies are not our true doppelgangers: they are psychologically different from us – not having phenomenal consciousness makes them unable to share all of our beliefs.⁴ So, the physicalist we are imagining may hold at the same time the following three claims: P entails C ; C does not apply to zombies (because they are not our true doppelgangers by not satisfying Ψ); and C accounts for E . How they defend the first claim is up to them, the point being that it does not imply the negation of the second.

It would also follow that zombies are either psychologically identical to us or lack phenomenal consciousness. If the latter, they do not satisfy Ψ (which holds for any true doppelgangers regardless of their physical constitution).⁵ If the former, they can have true beliefs self-attributing phenomenal states (and *eo ipso* phenomenal consciousness). To sum up, that P entails C does *not* in turn entail that C cannot explain E , because it is not the case that zombies satisfy P but not C (let us symbolize it as $[P(z) \ \& \ \neg C(z)]$). We can extend P as $(P \ \& \ \Psi)$ so that the first entailment becomes $([P \ \& \ \Psi] \rightarrow C)$. By the above argument, $(P[z] \ \& \ \neg \Psi[z])$, so zombies cannot be relevant for the truth or implications of $([P \ \& \ \Psi] \rightarrow C)$ – in particular, from the latter, $C(z)$ does not follow.

Not only can the second horn of the master argument's dilemma be called into question. According to Elpidorou (2013), the first horn poses no real problem to type B physicalists endorsing the PCS (any more than the problem posed by the explanatory gap).

The reason why P does not entail C is that the latter is not *a priori* derivable from the former because the latter (but not the former) essentially contains phenomenal terms. So, what lies behind the second horn is the well-known claim that phenomenal descriptions cannot be entailed by non-phenomenal descriptions (Elpidorou 2013: 1194) – the claim leading to an explanatory gap that type B physicalists acknowledge, do not intend to bridge, and whose purported ontological implications reject. (The same goes for the inference that if C explains E [so that C contains phenomenal terms], then P does not entail C). As

³ Note that if it is denied that the psychological supervenes on the physical, the boundary between hard and easy problems (Chalmers 1995) fades out (as all of them become hard ones [Chalmers 2006: 177]). The conceivability argument would have to rely on a much weaker intuition: zombies as beings that are physically/functionally identical to us but do not have the psychological processes we have (the intuition would probably be that they cannot be physically/functionally like us).

⁴ Note also that one does not have to be a physicalist to take for granted that C implies Ψ , or at least the part of Ψ having to do with phenomenal states, just as the right description of Ψ implies C . Thus, even for a non-physicalist, zombies might not satisfy C .

⁵ The physicalist we are imagining only adds that Ψ supervenes on P . Thus, that kind of physicalism might be thought as a way of denying that zombies – beings physically and psychologically identical to us but lacking phenomenal consciousness – are *secunda facie* conceivable.

the PCS acknowledges this kind of unbridgeable explanatory gap, it cannot be rejected for not having bridged it.

Elpidorou (2013: 1197) points out that, rather than derive C or Q *a priori* from P , what the PCS is supposed to do is account for any explanatory gap in a way compatible with physicalism – and that it does. The PCS contains a way for C to be compatible with physicalism despite C not being entailed by P (i.e., without the need to bridge the gap), namely, the thesis of conceptual isolation – the claim that, despite being logically independent, phenomenal and physical/functional concepts can refer to the same things.⁶

So, if $(P \ \& \ \neg C)$ is conceivable, then there is indeed an explanatory gap between P and C , but that is the same explanatory gap between P and Q – whose ontological consequences are said not to follow by type B physicalism. The gap between P and C does not jeopardize physicalism any more than that between P and Q , and through conceptual isolation, the PCS can explain why there are those explanatory gaps in a way compatible with physicalism.

To conclude, the master argument against the PCS has flaws that frustrate its aim. I pinpointed that the second horn depends on the claim that zombies do not satisfy C or E despite satisfying P and Ψ . However, we have seen that the anti-physicalist must admit that our epistemic situation with regard to consciousness is different from the zombie's because of a kind of mental state that zombies cannot have. In other words, the reason why zombies cannot satisfy E is that they do not satisfy Ψ . At the same time, we can think of the physicalist as holding that to satisfy C (and E), something must satisfy $(P \ \& \ \Psi)$. If so, zombies would be irrelevant for the truth of $([P \ \& \ \Psi] \rightarrow C)$. As to the first horn, it was noted that the new gap is the same gap which type-B physicalists admit, and that they are not supposed to bridge it but account for its rise consistently with physicalism – something they presumably do.

What Lies Ahead If We Follow Chalmers

Were we to concede the conceivability argument and the master argument, what options would be left? In this section, I want to analyze panpsychism, a view which Chalmers (2002, 2003a, 2015, 2017) has deemed a potential contender to an explanation of consciousness. I argue that its motivation is faulty and that it does not solve or explain what it was called to. Instead of aiming for comprehensiveness, my intention with this rather cursory detour is to show what lies ahead if we follow Chalmers's path.

Let us remember the argument from semantic stability, according to which P 's primary and secondary intensions might be identical. According to Chalmers, P 's primary and secondary intensions might differ only if different categorial properties (CPs) underlie in each world the physical property to which the phenomenon corresponds (Chalmers 2002: 187, 2010: 190).

⁶ Following Loar (1997), Elpidorou (2013: 17) considers that the implementation of conceptual isolation can be explained by the physicalist as a matter of conceptual role (see also Papineau 2006: 136-140).

Chalmers not only presupposes the distinction between dispositional properties (DPs, defined in terms of relationships to other properties) and CPs (are defined in terms of themselves), but that it creates two mutually exclusive and exhaustive sets of properties – where phenomenal properties are CPs and physical or functional properties are DPs. (Therefore, the only CPs we directly know about are phenomenal properties, while scientific knowledge informs us only about DPs). More precisely, phenomenal properties would arise above all from CPs – that could then be labeled *proto*-phenomenal. So, according to Chalmers (2002: 187-188), the only way in which the primary possibility of ($P \ \& \ \neg Q$) would not entail its secondary possibility is if panprotopsychism (the thesis that all CPs are thus protophenomenal) is true.

The prospects of panprotopsychism are, however, rather feeble. First, the distinction on which it depends not only has extravagant consequences but is not adequately motivated. Second, panprotopsychism is unable to do what it is supposed to do in the first place.

As already noted, Chalmers's argument requires the categorical/dispositional disjunction to be exclusive and exhaustive. However, such an exclusive disjunction has some implausible consequences: if the relation between causal roles and CPs is contingent (so that the latter can be related with different DPs), then it could happen that (unbeknownst to us) two different CPs could exchange the dispositional roles with which they are related, have the same roles at the same time, etc. (Orilia and Paolini 2022).

Heil (2003) has argued that the intuition of a contingent connection between CPs (or objects) and DPs may come from a confusion between predicates and properties – where a predicate (mistaken for a DP) is satisfied with different properties. He considers an alternative that concedes the contingency of the laws of nature but not a contingent relation between CPs and DPs, nor an exclusive disjunction between them: “*We are perhaps imagining worlds with different (though superficially similar) properties. Laws of nature would be contingent in so far as it is contingent that the actual world includes the properties it includes*” (2003: 94). Heil's (2003 Ch. 11) own preferred view on the CP/DP distinction is that every property is both dispositional and categorical (the ‘identity theory’; see also Martin 1993). According to the ‘identity theory’, ‘categorical’ and ‘dispositional’ would be two ways of partially considering one and the same property.

The second reason why panprotopsychism is hopeless has to do with its aptness to solve the hard problem of consciousness. Panprotopsychism opens its own explanatory gap (known as the ‘combination’ problem; James 1890: 160; Seager 1995), whereby it is vulnerable to the same kind of conceivability argument used to attack physicalism.

Protophenomenal properties are said to constitute phenomenal properties when ‘appropriately combined’, or when combined ‘in certain ways’ (Alter and Howell 2015: 283, 290). Not a word is said about what these ways are. Even worse, protophenomenal properties *cannot* explain phenomenal properties: insofar as they are defined as non-phenomenal, they are in the very same position as any other non-phenomenal property to explain consciousness. Because of that explanatory gap (or combination problem), a kind of zombie is conceivable that is microphysically and protophenomenally identical to one of us but lacks phenomenal properties (Goff 2009, 2017).

Chalmers (2017: 35) himself has recently admitted that panprotopsychism is implausible and that there seems to be no prospect of a solution for the combination problem. That seems to have plunged the position into a sort of nihilism: “No position on the mind – body problem is plausible. Materialism is implausible. Dualism is implausible. Idealism is implausible. Neutral monism is implausible. None-of-the-above is implausible” (Chalmers 2021: 29). All that pessimism is not mandatory, but it is certainly a consequence of granting the anti-physicalist arguments challenged in this manuscript.

Conclusions

The reason why zombies do not share our epistemic situation with regard to consciousness (E) is that they cannot have a kind of mental state that we have: our beliefs that truly self-attribute phenomenal states – these beliefs are inextricable from their causal history, which means that zombies cannot have these beliefs, since they cannot instantiate such causal chains. If, in order to satisfy C , something must satisfy Ψ (i.e., have all our psychological attributes), then it follows that zombies do not satisfy C . Physicalism would only add that nothing can satisfy P without satisfying Ψ .

As a consequence, the second horn of the Master Argument is objected: from $(P \rightarrow C)$ it does not follow that zombies satisfy C – because from $([P \& \Psi] \rightarrow C)$ it does not follow that zombies satisfy C . Zombies are simply irrelevant to the truth or implications of $(P \rightarrow C)$ (aka $([P \& \Psi] \rightarrow C)$). More generally, the Master Argument fails. The assumptions made in the first horn disarm it: type-B physicalism cannot be rejected because it does not *a priori* derive the phenomenal from the physical.

For its part, the conceivability argument in Chalmers’s version takes on water everywhere. ‘Primary intensions’ reflect possible meanings rather than actual meanings (and are hardly purely *a priori*). As for ideal conceivability, I argued that we cannot ascertain whether $(P \& \neg Q)$ falls within its scope. As for positive conceivability, I contended that we seem to lack a sufficiently rigorous notion of it. As a result, even though other forms of conceivability (*prima facie*, negative, and secondary) do not seem problematic, the argument we face (which depends on ideal, positive, and primary forms of conceivability) is not very compelling.

For years, conceivability arguments have been the spearhead of anti-physicalist views and have become a natural part of the landscape in debates on the metaphysics of consciousness. Nevertheless, the conceivability of zombies probably does not imply that they are metaphysically possible: the idea that it does (even in Chalmers’s most sophisticated version) is problematic, and there is a sound explanation for why it does not, which is compatible with physicalism (the PCS).

Finally, the inevitable consequence of granting what the conceivability and master arguments claim is a kind of theoretical nihilism. This is an effect of the realization that panprotopsychism is an unpromising option: it lacks a sound foothold (the view of the categorical/dispositional distinction as exclusive and exhaustive), and it is vulnerable to the explanatory gap.

Of course, it does not mean that physicalism or other forms of ontological monism are free of issues. On the contrary, the list goes on and on, including the exact nature of mind – brain relationships; the connection between psychological, phenomenal, and neural explanations; etc. Yet, if we are to move forward on the problem of consciousness, we must begin by leaving behind the idea that conceivability arguments have metaphysical consequences.

References

- Alter, T., Howell, R., 2015. The Short Slide from A posteriori Physicalism to Russellian Monism. In *Consciousness in the Physical World: Perspectives on Russellian Monism*, eds. T. Alter, Y. Nagasawa. Oxford: Oxford University Press, 277–299.
- Block, N., Stalnaker, R., 1998. Conceptual Analysis, Dualism, and the Explanatory Gap. *The Philosophical Review*, 108(1): 1–46.
- Chalmers, D., 1995. Facing up to the Problem of Consciousness. *Journal of Consciousness Studies*, 2(3): 200–219.
- Chalmers, D., 1996. *The Conscious Mind: in Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Chalmers, D., 2002. Does Conceivability Entail Possibility? In *Conceivability and Possibility*, eds. T. Gendler, J. Hawthorne. Oxford: Oxford University Press, 145–200.
- Chalmers, D., 2003a. Consciousness and its Place in Nature. In *The Blackwell Guide to Philosophy of Mind*, eds. S. Stich, T. Warfield. Hoboken: John Wiley & Sons, 102–142.
- Chalmers, D., 2003b. The Content and Epistemology of Phenomenal Belief. In *Consciousness: New Philosophical Perspectives*, eds. Q. Smith, A. Jokic. Oxford: Oxford University Press, 220–272.
- Chalmers, D., 2006. Phenomenal Concepts and the Explanatory Gap. In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, eds. T. Alter, S. Walter. Oxford: Oxford University Press, 167–194.
- Chalmers, D., 2010. *The Character of Consciousness*. Oxford: Oxford University Press.
- Chalmers, D., 2015. Panpsychism and Panprotopsyism. In *Consciousness in the Physical World: Perspectives on Russellian Monism*, eds. T. Alter, Y. Nagasawa. Oxford: Oxford University Press, 246–276.
- Chalmers, D., 2017. The Combination Problem for Panpsychism. In *Panpsychism: Contemporary Perspectives*, eds. G. Bruntrup, L. Jaskolla. Oxford: Oxford University Press, 179–214.
- Chalmers, D., 2021. Idealism and the Mind-Body Problem'. In *The Routledge Companion to Panpsychism*, ed. W. Seager. Oxford: Oxford University Press, 591–613.
- Elpidorou, A., 2013. Having it Both Ways: Consciousness, Unique not Otherworldly. *Philosophia*, 41: 1181–1203.
- Fraga, D., 2017. Ideal Reasoners don't Believe in Zombies. *Principia: An International Journal of Epistemology*, 21(1): 41–59.
- Goff, P., 2009. Why Panpsychism doesn't Help us Explain Consciousness. *Dialectica*, 63(3): 289–311.
- Goff, P., 2017. *Consciousness and Fundamental Reality*. Oxford: Oxford University Press.
- Heil, J., 2003. *From an Ontological Point of View*. Oxford: Oxford University Press.
- Hill, C., 1997. Imaginability, Conceivability, Possibility, and the Mind-Body Problem. *Philosophical Studies*, 87(1): 61–85.
- Hill, C., Mclaughlin, B., 1999. There are Fewer Things in Reality than are Dreamt of in Chalmers's Philosophy. *Philosophy and Phenomenological Research*, 59(2): 445–454.
- Jackson, F., 1982. Epiphenomenal Qualia. *Philosophical Quarterly*, 32: 127–136.
- James, W., [1890] 1981. *Principles of Psychology*, Vol. 1. Cambridge, MA: Harvard University Press.

- Kim, J., 1993. *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press.
- Loar, B., 1997. Phenomenal States. In *The Nature of Consciousness: Philosophical Debates*, eds. N. Block, O. Flanagan, G. Güzeldere. Cambridge, MA: MIT Press.
- Loar, B., 1999. David Chalmers's The Conscious Mind. *Philosophy and Phenomenological Research*, 59(2): 465–472.
- Martin, C. B., 1993. Power for Realists. In *Ontology, Causality, and Mind*, eds. J. Bacon, K. Campbell, L. Reinhardt. Cambridge: Cambridge University Press.
- Orilia, F., Paolini, M., 2022. Properties. In *The Stanford Encyclopedia of Philosophy*, ed. E. Zalta. Available at <https://plato.stanford.edu/archives/spr2022/entries/properties/> (Accessed 2 March 2023).
- Papineau, D., 2006. Phenomenal and Perceptual Concepts. In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, eds. T. Alter, S. Walter. Oxford: Oxford University Press, 111–144.
- Schroeter, L., 2004. The Rationalist Foundations of Chalmers' 2D Semantics. *Philosophical Studies*, 118: 227 – 255.
- Seager, W., 1995. Consciousness, Information, and Panpsychism. *Journal of Consciousness Studies*, 2(3): 272–288.
- Silins, N., 2021. Perceptual Experience and Perceptual Justification. In *The Stanford Encyclopedia of Philosophy*, ed. E. Zalta. Available at <https://plato.stanford.edu/archives/win2021/entries/perception-justification/> (Accessed 2 March 2023).
- Yablo, S., 1999. Concepts and Consciousness. *Philosophy and Phenomenological Research*, 59(2): 455–463.