Physicalism, dualism, and metaphysical gridlock

Katalin Balog
Rutgers Newark

The sacred .... doctrines! The very things that damn you
Each blessed time you stumble on them, since
You chop them so fine, sharpen them so neatly,
Refine them, twist them, that they drive you mad
Or turn to fetters. The human mind cannot
Sustain precise ideas - yet, in your pride,
It is precisely such that you are seeking... -
And you destroy yourselves in search of it.

Imre Madách, The Tragedy of Man Scene 7

Consciousness is frustratingly resistant to our efforts to understand it – especially where its relation to the physical world is concerned. During the last four decades, there has been a resurgence of arguments against physicalism and for varieties of metaphysical dualism about consciousness. The conclusion of these arguments is that phenomenal consciousness is absent from a world that is purely physical, that phenomenal consciousness involves fundamental, non-physical properties. While many contemporary philosophers of mind have found some of these arguments to be persuasive,² physicalists have reasonable rebuttals; they also have some

¹ I would like to thank Ned Block, David Chalmers, Troy Cross, Barry Loewer, Raymond Martin, Howard Robinson and Gilad Tanay, participants of a discussion group at NYU, audiences at the Rutgers University philosophy department, and at conferences at CEU and NYU for comments and helpful discussion of this paper.

² David Chalmers’ list of contemporary anti-physicalists includes Joseph Almog, Torin Alter, George Bealer, Laurence BonJour, Paul Boghossian, Tyler Burge, Tim Crane, John Foster, Brie Gertler, George Graham, W.D. Hart, Ted Honderich, Terry Horgan, Steven Horst, Jaegwon Kim, Saul Kripke, Harold Langsam, E.J. Lowe, Kirk Ludwig, Trenton Merricks, Martine Nida-Rümelin, Adam
arguments of their own against dualism. There is an impass between physicalism and anti-physicalism that seems hard to break. In this paper, I propose a novel way to look this impass: maybe the question of whether consciousness is purely physical does not have a determinate answer.

The anti-physicalist arguments start from a premise about a conceptual, epistemic, or explanatory gap between physical and phenomenal descriptions and conclude from this—a priori grounds—that if the phenomenal is real physicalism is false. I call these arguments “conceivability arguments”. Although not all of them have a premise concerning conceivability, they all, if successful, establish that it is inconceivable that phenomenal experience exists in a purely physical world. My first aim in this paper is to develop a master argument to counter these arguments. The master argument is crucially bolstered by what has come to be known as the phenomenal concept strategy; this strategy—following Brian Loar’s original proposal in (1997)—appeals to the special cognitive features of phenomenal concepts.


3 There are arguments in favor of physicalism that appeal to mental causation and the causal closure of physics (Loewer 1995), (Papineau 1995). Loewer and Papineau argue that the anti-physicalist is forced into adopting one of these implausible positions: epiphenomenalism, causal overdetermination, or denial of the causal closure of physics.

4 Phenomenal descriptions attribute phenomenal properties to experience (and perhaps to thought) in the sense of there being something it is like to undergo an experience, something one can normally introspect, e.g., the feeling of my toes flexing that (partly) characterizes my present bodily sensation. I will assume throughout the paper that there are phenomenal properties in this sense. For eliminativism about phenomenal properties, see, e.g., Rey (2007) and Frankish (2016).


6 I will later introduce and explicate two notions of conceivability.

7 Stoljar’s (2005) phrase.
in providing a physicalistically respectable explanation of the various gaps.

In the second part of the paper I assess the dialectical situation involving the conceivability arguments and the master argument against them and argue that there is a puzzling symmetry between dualist attacks on physicalism and physicalist replies. Each position can be developed in a way to defend itself from attacks from the other position. In the end, there are neither a priori nor a posteriori ways to decide between the two; the question concerning the nature of consciousness seems unanswerable. I conclude the paper by giving some reasons to think that the question of whether physicalism or anti-physicalism is true is unanswerable because it does not have a determinate answer.

The plan is as follows. Section I provides the metaphysical background. Section II and III discusses David Chalmers’ zombie argument and the physicalist master argument that rebuts all versions of the conceivability argument. In section IV, I discuss how the phenomenal concept strategy complements and supports the counter conceivability argument. In section V, I argue that physicalists and anti-physicalists are in a gridlock: they are both able to defend themselves from the attacks of the other side, and at the same time they both can be viewed as question-begging from the other side. Finally, in section VI, I propose a novel approach to the mind-body problem that explains the gridlock; I argue that perhaps there is no determinate fact of the matter as to the nature of consciousness.

I. Metaphysical background

The debate between physicalism and anti-physicalism is about fundamental ontology.

a) Physicalism. According to physicalism, the world’s fundamental ontology is physical.\(^8\) Intentionality and consciousness is instantiated in macroscopic systems in virtue of immensely complex arrangements of fundamental physical properties and entities and their causal/nomological features; i.e. for biological individuals in virtue of brain states and

\(^8\) Contemporary physicalists typically hold that the best account of that ontology is provided by fundamental physics. Physics’ best hypotheses about fundamental ontology is that it consists of elementary particles, strings and/or fields occupying a space-time structure, and possessing a limited number of quantitative properties (mass, charge, electromagnetic potential, and so on). Physics also claims that there are only a few fundamental dynamical and perhaps non-dynamical laws that govern the structure of space-time and the evolution of its occupants.
processes.

Following Frank Jackson (1993), I will assume that there is a fundamental vocabulary (although not necessarily in our language as it is currently) in which there is a complete fundamental true description of the world. This description specifies the total spatio-temporal distribution of fundamental entities, the totality of instantiations of fundamental properties and relations, and the fundamental laws. Furthermore, I will stipulate that if physicalism is true then none of the elementary vocabulary refers to mental or proto-mental entities or properties.

Jackson pointed out that a necessary condition for the truth of physicalism in a world is that all positive truths in that world, including, as the case may be, positive truths about phenomenal consciousness,⁹ are metaphysically necessitated by the complete physical truth about that world.¹⁰ Accordingly, if physicalism is true in our world, the Physicalist Entailment Thesis is true as well:

\[(\text{Phys}) \forall T (\Box (P \Rightarrow T)).^{11}\]

---

⁹ A positive phenomenal statement says that a phenomenal property is instantiated; e.g., *Joe is feeling an itch*. Negative truths, like *There are no angels*, and global statements, like *Every gold cube has a volume smaller that one cubic meter*, are not metaphysically necessitated by the complete physical truth about the world P although they are necessitated by P and a clause that says that P is the whole fundamental truth. However, the phenomenal and physical truths we will be interested in are all positive truths so from here on I will ignore this complication.

¹⁰ This formulation is based on Jackson’s (Jackson 1993). The first precise formulation of physicalism along these lines is due to Lewis (Lewis 1983a). Subsequent discussions are variations on the same theme. Many philosophers, among them non-physicalists, accept this formulation as capturing a very important component of the intuitive idea of physicalism. But it doesn’t express the full physicalist commitment – only a necessary condition – because it is apparently compatible with certain ontologies that are intuitively non-physicalist e.g., with one in which there are fundamental mental as well as fundamental physical properties connected by “brute” necessary connections.

¹¹ \(\forall\) is a substitutional quantifier, \(T\) is a statement variable for true positive statements, \(\Box\) is the metaphysical necessity operator, and \(P\) is the complete fundamental physical description of the world, including the fundamental physical laws.
It follows that if there are mental truths – for example, that Mary knows what it is like to see red – that are not necessitated by the complete physical description then physicalism is false.

b) Anti-physicalism. According to anti-physicalism, the fundamental ontology of the world is not exhausted by the physical. Anti-physicalism comes in a number of different varieties. The usual suspects include non-interactive and interactive property and substance dualism, panpsychism, and idealism. In the end, I think only two of these are serious contenders. I find idealism too implausible to take seriously; and substance dualism strikes me as positing a metaphysical entity (a simple substance or soul) that we have no good reason for believing in.\footnote{However, see, e.g., Nida-Rümelin 2013 for an argument for mental substance.} By “interactive property dualism” – for lack of a better word – I mean property dualism that is committed to a denial of the causal closure of physics, i.e., committed to the idea that certain phenomena – most likely, purposeful behavior – does not have a fully physical explanation. Such a view could, in principle, be empirically verified. I don’t suppose that evidence in its favor cannot turn up. But such evidence hasn’t turned up so far and the prospects for it are not promising; for the purposes of this paper I assume interactive dualism is false. This leaves non-interactive property dualism (which is somewhat of a misnomer as it just means dualism that also accepts the causal closure of physics – not necessarily a view denying mental causation) and panpsychism as the strongest candidates.

Non-interactive dualism holds that the physical does not necessitate all truth, and that the fundamental ontology of the world includes phenomenal and/or proto-phenomenal properties. It also posits the existence of fundamental vertical laws\footnote{I will assume that these laws are contingent; i.e., not metaphysically necessary. If laws are taken to be metaphysically necessary then it is difficult to state the difference between physicalism and dualism since then both would hold that configurations of physical property instantiations metaphysically necessitate mental property instantiations.} that connect arrangements of physical entities and properties to phenomenal or protophenomenal properties. Furthermore, non-interactive dualism – which, to be sure, is not a denial of mental causation per se – is committed to the causal closure of physics. Consequently, it is committed either to epiphenomenalism or causal overdetermination of the physical by the mental.

Panpsychism is back in vogue due to the frustration over the impasse between physicalism
and property dualism.\textsuperscript{14} Its most plausible version, Russellian panpsychism – a.k.a. Russellian monism or dual aspect monism – holds that the intrinsic, as opposed to causal/dispositional, nature of fundamental properties and entities is phenomenal or proto-phenomenal.\textsuperscript{15} It too is committed to the causal closure of physics since it holds that the causal dispositional nature of fundamental properties is physical. Russellian panpsychism allows for mental causation; according to it, the intrinsic nature of the causally relevant physical properties involved is itself mental. The central consideration in its favor is the claim that physics as it is is incomplete; it doesn’t tell us what physical things are like \textit{in themselves}, intrinsically, so to speak. This is a controversial claim, of course, and even if it weren’t, Russellian panpsychism still would face the “combination problem”, i.e., the problem of explaining how micro-experiences give rise to macro-experiences. It is hard to see how the experiences of micro-entities could combine to give rise to a composite experience had by a composite entity. It is equally hard to see how (non-phenomenal) proto-phenomenal properties could combine to yield phenomenal properties.

It is sometimes argued that Russellian panpsychism is compatible – at least with the letter of – physicalism: after all, it says that the arrangements of fundamental physical entities and their physical properties is metaphysically sufficient for the instantiation of phenomenal properties, as long as the intrinsic natures of fundamental entities and properties are taken into account. Stoljar (2001), e.g., considers Russellian panpsychism a version of physicalism. Nothing important hangs on the terminology; however, I think it is useful to firmly distinguish Russellian panpsychism from the version of physicalism I will be mostly concerned with in this paper, according to which physicalism requires that \textit{fundamental} physical properties and entities are non-mental and not even proto-mental. On this definition, physicalism and Russellian panpsychism are incompatible.

If Russellian panpsychism is true both physicalism and non-interactive property dualism are false. On the other hand, as I will assume in this paper, Russellian panpsychism is false, the two views left standing are physicalism and non-interactive property dualism. These are the views I will consider in the rest of this paper.

\textsuperscript{14} See, e.g., Strawson 2006 and Chalmers 2017.

\textsuperscript{15} Proto-phenomenal properties are not full-blown phenomenal properties; but they are constituents of full-blown phenomenal properties.
II. The Zombie Argument and the inconceivability of illuminati

There is a line of argument against physicalism that goes back at least to Descartes’ argument for the distinctness of mind and body. In fact, these arguments can be understood to conclude, on the basis of a priori considerations, that phenomenal properties cannot be exemplified in a purely physical world. The descendent of this argument that has received the most attention in the last decade is David Chalmers’ “Zombie Argument.”

The Zombie Argument

1) $P \& \neg Q$ is conceivable.  

2) If $P \& \neg Q$ is conceivable then $P \& \neg Q$ is metaphysically possible (CP principle)

3) If $P \& \neg Q$ is metaphysically possible then physicalism is false.

4) Physicalism is false.

Chalmers (2009) suggests that the relevant notion of conceivability that figures in the Zombie Argument is what he calls “positive conceivability”\(^{18}\), which involves being able to form some sort of clear and distinct conception of a situation in which the hypothesis is true.\(^ {19}\)

The Zombie Argument is valid. Premise 3 follows from the definition of physicalism. Philosophers who think that there is a functional or representational analysis of phenomenal

---

\(^{16}\) Zombies are creatures that are physically identical to normal humans but have no phenomenal experiences whatsoever.

\(^{17}\) $P$ is the complete fundamental physical description of the world, including the fundamental physical laws, and $Q$ is a positive phenomenal truth, e.g., that someone is having a visual experience with a particular phenomenal character at a particular time.

\(^{18}\) Chalmers introduces a battery of different conceivability concepts in (2002a), among them negative conceivability, positive conceivability, ideal conceivability, etc.

\(^{19}\) Running the argument based on the notion of positive conceivability renders it immune to the Zombie Refutation I introduced in Balog (1999). Hence the need for the physicalist to come up with another counterargument; I propose one later in the paper.
consciousness reject premise 1.\textsuperscript{20} But I agree with Chalmers that no physical description \textit{a priori} entails any positive phenomenal description. Later I will offer some considerations based on the nature of phenomenal concepts for why this is so but for now I will just assume that premise 1 is true. So for both Chalmers and myself the crucial premise in the argument is 2.\textsuperscript{21}

The conceivability arguments are based on a premise concerning the epistemic status of phenomenal experience with respect to the physical (e.g., that zombies are \textit{prima facie} possible (Kripke 1972); that zombies are conceivable (Chalmers 1996, 2009); that there is an explanatory gap (Levine 2001); that our grasp of phenomenal properties is unconnected to our grasp of physical properties (Nida-Rümelin 2007); or that phenomenal and physical concepts are independent yet connotate essential properties of the referent (White 2007)), and another premise linking this epistemic status to metaphysical status. Both kinds of premise, if true, are a priori true.\textsuperscript{22} The conceivability arguments all conclude that physicalism is false. Arguably, this conclusion is not a priori knowable, even according to dualists, since it depends on the – plausibly a posteriori – assumption that phenomenal experience exists.\textsuperscript{23} However, I propose

\textsuperscript{20} E.g. Lewis 1966 and Jackson 2003. Analytic functionalism or representationalism concerns the \textit{meaning} of phenomenal terms; it says that such meanings can be analyzed in functional or representational terms. \textit{Pain}, e.g., according the analytic functionalism, has a conceptual role that connects it (in the meaning-constituting way) with complex concepts like \textit{typically caused by injury}, \textit{typically causes avoidance behavior}, \textit{typically cases saying “ouch”}, etc. Analytic functionalism or representationalism rebuts the conceivability arguments by denying the conceptual, epistemic, and explanatory gaps between physical and phenomenal descriptions. Analytic functionalism/representationalism, of course, has to explain why there seem to be such epistemic gaps when in reality there aren’t. See also Kirk (2005) for an interesting argument against the conceivability of zombies whose grounds go beyond analytic functionalism.

\textsuperscript{21} Chalmers’ defense and development of the two-dimensional framework and of the conceivability-possibility link can be found in Chalmers and Jackson 2001, Chalmers 2002a, and Chalmers 2004. There are important discussions in Block and Stalnaker (1999), Yablo (1993, 2002), and Soames (2004). I briefly discuss what I think goes wrong with a related argument by Frank Jackson (Balog 2001).

\textsuperscript{22} One might question the a priori status of either of these premises; however, the proponents of these arguments do all seem to regard their premises as a priori so I am not going to pursue this complication here.

\textsuperscript{23} Perhaps one can argue to the contrary, claiming that there is a sense in which we can know a priori in the first person that we have phenomenal experience - but it is not quite the same sense as we know
that each of the dualist conceivability arguments is only as good another one that wears the a priority of its premises on its sleeves. One can tweak the formulation of the arguments – bringing in nothing the dualist couldn’t accept – to make explicit their a priori elements. It follows that if these arguments are successful then what I call “illuminati” – purely physical creatures that are our physical duplicates and enjoy phenomenal experiences – are inconceivable. The debate between the anti-physicalist and the physicalist then comes to this: the anti-physicalist denies, and the physicalist asserts the conceivability of illuminati.

The Counter Conceivability Argument is a master argument that provides a rebuttal of all the conceivability arguments via an argument for the conceivability of illuminati. In what follows, I again consider the particular case of the Zombie Argument.

*The Zombie Argument:*

1. $P \& \sim Q$ is conceivable.
2. If $P \& \sim Q$ is conceivable then $P \& \sim Q$ is metaphysically possible. (CP principle)
3. If $P \& \sim Q$ is metaphysically possible then physicalism is false.

24 Note that “purely physical” is not meant to exclude the presence of mental properties; only the presence of fundamental mental properties; so illuminati are not trivially inconceivable.

25 Physicalists, of course, are committed to the actual existence of illuminati. But the argument will only require their conceivability.

26 It can be shown that if the argumentative strategy works against the Zombie Argument, it works against the other conceivability arguments as well – but I will not spell out the details. The basic idea is that these arguments can all be formulated to employ supposedly a priori premises for the (supposedly a priori) lemma that if phenomenal experience exists then physicalism is false.

27 $P$, again, is the full physical description of the world, including the fundamental physical laws, and $Q$ is a positive phenomenal truth.
4) Physicalism is false.

As I have said, it can be shown that – on plausible assumptions about the a priori status of the Zombie Argument’s premises – one can deduce from this the inconceivability of illuminati:

ILLUMINATI-INCONCEIVABLE: It is a priori that if Q is true then physicalism is false.  

The actual deduction is somewhat technical. Interested readers can find it in the Appendix.

III. The Counter Conceivability Argument

There is a case to be made – though I won’t make it here – that the inconceivability of illuminati follows not just from the Zombie Argument, but from all other conceivability arguments as well. Consequently, one can rebut them all by offering reasons for the conceivability of illuminati.

It is worth noting that the notion of conceivability involved here is weaker than the notion of positive conceivability invoked in the Zombie Argument. All that the physicalist needs to show is that illuminati are conceivable in what Chalmers calls the negative sense, that is, that their existence cannot be ruled out on a priori grounds. So to rebut the conceivability arguments, the physicalist only needs to argue for the negative – as opposed to positive – conceivability of illuminati.

But how can the physicalist argue for the negative conceivability of illuminati? It will be

28 Marton (1998) expresses a similar idea to the effect that the Zombie argument can be understood to imply the inconceivability of the supervenience of the phenomenal on the physical.

29 Positive conceivability of a statement requires some sort of clear and distinct conception of a situation where the statement is true.

30 This is true in spite of the fact that some of the supposedly a priori truth involved in ruling out the existence of illuminati employ the notion of positive, rather than negative conceivability.

31 The physicalist might want to argue for the positive conceivability of illuminati as well; if illuminati were positively conceivable that would amount to an knockdown refutation of the Zombie Argument. As we will see later, the Counter Conceivability Argument doesn’t amount to such a knockdown refutation – but it is but this is not strictly necessary for the success of the Counter Conceivability Argument.
instructive to compare what is involved in the negative conceivability of zombies as opposed to the negative conceivability of illuminati. One can see that zombies are negatively conceivable immediately, without having to consider anything else, e.g., facts about how the brain – or anything else in the physical world – works. Phenomenal concepts refer to phenomenal experiences in a substantial yet direct way, i.e., they do not have analytically sufficient conditions in physical, functional, or behavioral terms; this results in enough conceptual independence of phenomenal and physical concepts to make zombies negatively conceivable. How about illuminati? The basic idea is this. Zombies are negatively conceivable because we grasp phenomenal properties in a substantial yet direct way; but illuminati also appear negatively conceivable, and for the very same reasons. There is no conceptual contradiction in the claim that every instantiated phenomenal property is identical with some physical or (physically realized) functional property or other. Such identifications do not seem to be incoherent. Perhaps even stars or rocks are conceivably conscious in virtue of their physical properties. One might object to this on the grounds that only entities with some minimal cognitive organization can be conceived to have phenomenal states. While no functional or physical description is analytically sufficient for a state to be phenomenal perhaps there are some analytically necessary conditions for phenomenality, requiring that subjects have at least a rudimentary cognitive structure. If these analytically necessary conditions exist, they plausibly rule out stars and rocks as bearers of phenomenal experiences. They would rule out panpsychism as well. Whether or not this is right, there is certainly no conceptual contradiction in identifying qualia with physical states playing whatever functional roles are necessary to satisfy those conditions.

Strictly speaking, conceiving of a world where such property identities hold is not thereby conceiving of a world where physicalism is true. Perhaps there are metaphysically possible worlds where physicalism is false for reasons having nothing to do with phenomenal properties (e.g. because of the instantiation of some non-phenomenal non-physical property). However, I don’t see any reason why, in conceiving of a world where psycho-physical

---

32 Levine (2007) and Chalmers (2007) both talk about the “substantiality” of our conception of phenomenal states and of our phenomenal knowledge. Later I will explain in more detail what I mean by “substantiality”. The main idea is that we grasp phenomenal states in a way that seems to reveal their essence.

33 At least in the phenomenal way of grasping them; if physicalism is right, then, of course we can grasp them in a physical way as well. This doesn’t affect the first point, though.
identities hold, one cannot just stipulate that there are no further non-physical, or non-physically realized properties. In any case, the conceivability arguments purport to rule out physicalism by ruling it out on a priori grounds that physical facts necessitate phenomenal facts, so arguing that such necessitation is conceivable answers those arguments.

So far so good. But there is a problem. Even if – as I argued – illuminati cannot be ruled out on conceptual grounds, perhaps they can be ruled out via the putatively a priori CP Principle? It seems question-begging for the physicalist to rest their case about the conceivability of illuminati merely on conceptual coherence. Physicalists need to do more; they need to respond in some stronger fashion to the CP Principle. Showing, e.g., that the CP Principle leads to contradiction would be the most effective response to the dualist. One way to do that would be to show that illuminati are not only negatively conceivable, but also positively conceivable – which is the relevant notion of appealed to in the CP Principle. In effect, there is reason to think that the positive conceivability of zombies and the positive conceivability of illuminati are on a par. Both are equally prima facie conceivable, due precisely to the direct and substantial grasp of phenomenal properties that phenomenal concepts afford us. Moreover, and this is the key, the dualist cannot use the CP Principle against the positive conceivability of illuminati (in the way she might try to use it against the negative conceivability of illuminati): if the dualist argues that the positive conceivability of illuminati can be ruled out a priori on the basis of the positive conceivability of zombies and the CP Principle, the physicalist can counter that this is just special pleading. After all, there is an argument of the exact same form, using the positive conceivability of illuminati and the CP Principle to rule out the positive conceivability of zombies.

---

34 This is what I have done in an earlier paper (Balog 1999), with respect to the CP Principle, understood as involving mere negative (and not positive) conceivability. I have argued that, on some plausible assumptions, the principle undermines itself by leading to a contradiction.

35 Frankish (2007) attacks the Zombie Argument on the grounds that “anti-zombies” are conceivable. Though he doesn’t discuss the distinction between positive and negative conceivability, I think his claim that “anti-zombies” are conceivable is best interpreted as the positive conceivability of illuminati. Sturgeon (2000) and Brown (2010) also gives an argument similar in form to show that conceivability cuts both ways and so that the Zombie Argument doesn’t have force. These both differ from my Counter Conceivability Argument. Piccinini (2015) uses considerations from modal logic to find symmetry between conceivability arguments for and against dualism.

36 An argument with a similar structure is suggested in Loewer (1978) with respect to Leibniz’s Ontological Argument.
Though this is an interesting argument I will sidestep it for the purposes of this paper. For one thing, it applies only to the Zombie Argument and not to the other anti-physicalist arguments; but there are also questions about the positive conceivable of illuminati – that is, whether we have a clear and distinct conception of an illuminata world – which I do not want to pursue here. So what if a knock-down argument is not possible with regard to the CP Principle? How is the physicalist going to deal with the Zombie Argument? In particular, how does the focus on the negative conceivable of illuminati help? So far what we have seen is that the physicalist can show that illuminati are negatively conceivable modulo the CP Principle. But the CP Principle still needs to be answered. Here is the idea. The physicalist might not be able to outright refute the CP Principle. She might, however, be able to extend her argument for the conceivable of illuminati by not only considering the conceptual coherence of psycho-physical identities, but considering it in the context of a holistic account of laws, modality, mental causation, concepts in general and phenomenal concepts in particular, all of which fits together to provide – among other things – a physicalist account of the epistemic gaps between phenomenal and physical descriptions from which the anti-physicalist arguments take off.\(^{37}\) If such an account is possible it would count as an argument against the CP Principle and so for the negative conceivable of illuminati.

A look at the CP Principle itself supports this diagnosis. Where does the supposed a priori status of the CP Principle come from? It certainly doesn’t seem to be a conceptual truth. Besides citing metaphysical intuition, the dualist has other options as well. Chalmers (2002), e.g. argues that the CP Principle provides the simplest and explanatorily most satisfying account of modality. It also best explains the puzzling epistemic gaps the conceivable arguments exploit. Even if it did – which the physicalist surely questions – this makes the CP Principle importantly different from stock examples of a priori truths, e.g., \(2+2=4\), or \textit{Bachelors are unmarried}, whose falsity can be ruled out by a priori methods independent of holistic considerations about how well one’s overall metaphysical framework accounts for modality, mental causation, laws, meaning, etc. The a priori status of the CP Principle, in contrast, depends on these same holistic considerations.\(^{38}\) Consequently the Zombie Argument, based on the CP Principle, cannot demand assent from all parties to the debate – like valid arguments based on premises like \(2+2=4\), or \textit{Bachelors are unmarried} can –

\(^{37}\) Hill and McLaughlin (1999) makes a similar proposal in the context of an argument for physicalism.

\(^{38}\) I am not going to decide whether the terminology \textit{a priori} properly describes such considerations. Nothing rides on the terminology.
independent of their metaphysical outlook. There is a physicalist argument to be made on these general grounds against the apriority of the CP Principle, and for the conceivability of illuminati. In the rest of the paper I will follow out this dialectic in broad strokes.

IV. The Phenomenal Concept Strategy

As part of an overall physicalist strategy to answer the anti-physicalist arguments, I will now invoke an approach to the mind-body problem which has become known as the “phenomenal concept strategy”. The idea is to think about our epistemic/conceptual relation to consciousness – the conceivability of zombies, the explanatory gap, our substantial grasp of phenomenal experiences, etc. – in terms of the peculiar nature of phenomenal concepts, rather than in terms of the peculiar nature of phenomenal experience itself. The key factor is that the explanation on offer is compatible with physicalism.

The core idea

The physicalist claim that illuminati are conceivable can be boosted by a physicalistic explanation of how one can be aware of phenomenal states in a direct and substantial way. Such an account would simultaneously support physicalism and undercut the rationale behind the CP Principle. I think we have such an account: the constitutional account of phenomenal states. On this account, there is an intimate relation between phenomenal concepts and their referents; token experiences serve as modes of presentation of the phenomenal properties.

39 I am indebted to Troy Cross for discussion on this issue.

40 An argument of this sort for the conceivability of illuminati is at the same time an argument for the actual truth of physicalism. Such a holistic argument for the coherence of physicalism – which removes the main reason against it – counts as reason to believe in the actual truth of physicalism.

41 I defend a version of that account in Balog 2012a. Other versions of the constitutional account have been proposed, on the physicalist side, by Hill and McLaughlin (1999), Block (2007), and Papineau (2002, 2007); David Chalmers (2003) also put forward a version of this account on the dualist side.

42 There are other proposals that fall under the general umbrella of the phenomenal concept strategy. Not all of these proposals acknowledge the substantial manner in which we are aware of our own phenomenal states. They include recognitional accounts (Tye 2003), demonstrative accounts (Levine 2007, Perry 2001), and information-theoretical accounts (Aydede & Güzeldere 2005).
they instantiate.\textsuperscript{43} In the case of most concepts, e.g., the concept WATER, it doesn’t matter exactly what neural configurations constitute a particular token of WATER as long as the requisite causal/informational relations between it and water hold. But in the case of phenomenal concepts, e.g., the concept PAIN, constitution matters for reference, both in terms of how the reference is determined, and in terms of how the concept cognitively presents its reference. More precisely, on this view, every token of a phenomenal concept applied to current experience is constituted by \textit{that very} token experience, and this fact is crucial in determining the reference of the concept. Not only is it the case that a token experience that realizes a token concept instantiates a token of the referent of the concept, but it is because the concept is so constituted that it so refers. There are, of course, applications of phenomenal concepts that are, on this theory, \textit{not} constituted by token experiences; e.g., applications of phenomenal concepts to one’s past or future experience, to other peoples’ experiences, etc. But the canonical, first person, present tense applications are always so constituted and the other applications are dependent on the first person applications.\textsuperscript{44}

This account explains the puzzling aspects of our epistemic/conceptual relation to our conscious states in a manner consistent with physicalism. First of all, the constitutional account explains how we can have a \textit{substantial grasp} of phenomenal properties even while this grasp is direct, and unmediated by physical or functional modes of presentation. Because, on this account, in the canonical, first person, present tense applications of a phenomenal concept an instance of the property it refers to is \textit{literally} (physically) present in the concept, there will be always something it is like to token the concept in those canonical applications.\textsuperscript{45} Undergoing a token of the phenomenal property reveals something essential about that property, namely, it reveals \textit{what it is like to have it}. This means that phenomenal concepts provide a substantial grasp of the phenomenal properties they refer to. And because, according to the theory, tokens of phenomenal concepts present their referent as the property

\textsuperscript{43} I consider experiences as mental states that have their phenomenal properties essentially, but instantiate other (physical, functional) properties. Also, in what follows I concentrate on phenomenal concepts that refer to phenomenal \textit{properties}; but the account can be easily modified to apply to concepts that refer to \textit{particular instances} of phenomenal concepts.

\textsuperscript{44} Such “indirect” applications of phenomenal concepts stand in an intricate conceptual relation with the “direct”, first person present tense applications. For an account, see (Balog 2012a).

\textsuperscript{45} Levine (2006, 2007) is critical of this approach. He argues that it is impossible to explain \textit{cognitive presence} by \textit{physical presence}. 15
whose token they incorporate – and not via any functional or physical description – they will refer to phenomenal properties directly, as well as substantively. The constitutional account can now be marshaled to account for the epistemic/conceptual gaps that drive the conceivability arguments.\footnote{For a detailed account of how this works see Balog 2012a.}

The conceivability of zombies is explained by the directness and substantiality of our direct phenomenal concepts which, under the constitutional account, is compatible with physicalism. The directness of phenomenal concepts follows from the fact that the reference of a direct phenomenal concept is determined by how it is constituted and \textit{not} by any description that is associated \textit{a priori} with the concept. Phenomenal concepts are supposed to be different in this way from concepts like WATER and even name concepts like CICERO. Chalmers and Jackson (2001) claim that these concepts are associated \textit{a priori} with descriptions (e.g. “the transparent potable liquid...”, “the Roman orator who is at the origin of a causal chain culminating in this token”) and these connections are \textit{sufficient} to obtain \textit{a priori} entailments from the full fundamental description to all positive statements of fact, e.g., they are sufficient to rule out \textit{a priori} a scenario where everything is physically the same but yet there is no water. In other words, a situation where everything is physically the same but yet there is no water is inconceivable. One doesn’t have to commit to this to see that zombies are conceivable; however, the conceivability of zombies can only have significance if this is the case.\footnote{This is because the conceivability of zombies only has significance if one also accepts the CP Principle – which one clearly couldn’t do if it was conceivable that everything is as it is physically but there is no water as it would render the CP Principle dead on arrival. For an argument for the equivalence of the \textit{a priori entailment thesis} and Chalmers’s Master Principle that lies behind the CP Principle, see my doctoral thesis (Balog 1998), p 124-5.} I am going along with the idea of \textit{a priori} entailment – except, of course, for phenomenal statements – for the sake of argument; Chalmers’ argument can be rebutted in a way that doesn’t depend on rejecting its semantic framework wholesale.\footnote{But see, e.g., Block and Stalnaker (1999) and McLaughlin (2007) for an argument that these entailments are not \textit{a priori} even for positive non-phenomenal statements.}

The point is that even if one accepts that all true non-phenomenal statements are \textit{a priori} entailed by the full fundamental description of the world, one can still claim that it is not so with respect to phenomenal concepts. Because of the fundamentally different cognitive

\begin{footnotesize}
\begin{enumerate}
\item For a detailed account of how this works see Balog 2012a.
\item This is because the conceivability of zombies only has significance if one also accepts the CP Principle – which one clearly couldn’t do if it was conceivable that everything is as it is physically but there is no water as it would render the CP Principle dead on arrival. For an argument for the equivalence of the \textit{a priori entailment thesis} and Chalmers’s Master Principle that lies behind the CP Principle, see my doctoral thesis (Balog 1998), p 124-5.
\item But see, e.g., Block and Stalnaker (1999) and McLaughlin (2007) for an argument that these entailments are not \textit{a priori} even for positive non-phenomenal statements.
\end{enumerate}
\end{footnotesize}
architecture of phenomenal concepts, there are no a priori connections between phenomenal and physical/functional/structural concepts that are sufficient to rule out a priori the zombie scenario – and this is perfectly compatible with physicalism. *Nota bene:* I am not denying that there are inferential links between thoughts involving direct phenomenal concepts that are individuative of them. I think it is quite plausible that there are conceptual links, even perhaps concept individuative conceptual links between direct phenomenal concepts such as we apply our own occurrent phenomenal experience on the one hand, and other direct phenomenal concepts, or even indirect phenomenal concepts such as we apply to other people’s phenomenal experiences on the other. My point is that to the extent that these are *a priori* they do not add up to conceptually sufficient conditions in terms of other mental concepts, functional, or behavioral concepts, etc. in other words, they are not of the sort that enables one to rule out *a priori* the zombie-scenario. So the point remains that because phenomenal concepts are direct and substantial, we can conceive of zombies. This explanation is perfectly compatible with a physicalist – as well as a dualist – metaphysics and leaves the CP Principle in place for all except phenomenal statements.

The constitutional account can also explain other puzzling features of our epistemic relation to phenomenal experience, like the explanatory gap, the incorrigibility of certain of our phenomenal judgements, the semantic stability of phenomenal concepts, etc. Of course, for the constitutional account to work, some of its details have to be worked out, in particular, the idea that constitution plays a role in determining reference in the case of phenomenal concepts. Whether this can be fully done still remains to be seen.

V. Metaphysical gridlock

Let’s consider again the dialectic between the physicalist and dualist proponents of the conceivability arguments. The phenomenal concept strategy notwithstanding, it must be granted that the dualist can still hold on to the claim that, as opposed to zombies, illuminati are not conceivable after all given the CP Principle. However, the phenomenal concept strategy does help to *undermine* the intuitive strength of the CP Principle. It explains how the CP Principle could appear plausible even in a purely physical world, thereby undercutting the

---

49 For a more detailed account of all of these, see Balog (2012a).

50 For a proposal see Balog (2012a). Chalmers (2007) argues that the phenomenal concept strategy cannot work for very general a priori reasons. For a reply see Carruthers and Veillet (2007), and Balog (2012b).
rationale behind it. Earlier I indicated that the argument for the conceivability of illuminati is a holistic affair so the physicalist also needs to present an alternative to the CP Principle, and to dualist metaphysics in general, by presenting an account of fundamental properties, laws, concepts, modality, etc. compatible with physicalism. Much can and has been said about these issues and I won’t attempt to adjudicate it here.\(^{51}\)

Rather, I will assess the dialectic between the physicalist and the anti-physicalist on the assumption that the general theoretical-metaphysical considerations are not decisive in either direction. If this is the case, then the issue has been fought to a draw. By her own lights, the physicalist has rebutted the anti-physicalist arguments. From her point of view, the CP Principle is demonstrably mistaken. However, the dualist is in a similar situation. According to the dualist, for all the clever verbal feats of the phenomenal concept strategy, illuminati are still inconceivable, as long as one is holding tight to the CP Principle. Let’s pause here for a moment. On the one hand, the anti-physicalist argues that the CP Principle is a priori true, and uses it to show that the physicalist efforts to undermine them are unsuccessful. On the other hand, the physicalist argues that illuminati are conceivable – based on the phenomenal concept strategy and holistic considerations about metaphysics – and uses this claim to undermine the CP Principle. By their own lights both sides seem justified in holding on to their key doctrines and denying those of their opponents. Where you end up depends on what you take as your starting point. And, as far as I can see, neither side has a privileged start.

This is puzzling. One would have thought that when it comes to a priorities like the CP Principle, or the conceivability of bright illuminati, there are a priori ways to justify or refute them. But it seems like there are no principles outside the physicalist and anti-physicalist systems that could settle this issue. What we have here is a puzzling symmetry between the two positions. The situation appears to be a stalemate.

\(^{51}\) On the physicalist side, there are arguments in favor of physicalism that appeal to mental causation and the causal closure of physics (Loewer 1995), (Papineau 1995). Loewer and Papineau argue that the anti-physicalist is forced into adopting one of these implausible positions: epiphenomenalism, causal overdetermination, or denial of the causal closure of physics. Hill and McLaughlin (1999), and McLaughlin (2007) argue more generally that there are powerful reasons to prefer physicalism as an overall explanatory metaphysics over dualism. See also Melnyk (2003) for discussion of a history of successful reduction of higher level properties to lower level ones. On the dualist side, Chalmers, e.g., argues (2002a, 2002b) that the CP Principle – and with it, dualism – is required to give a satisfying account of modality.
It is unlikely that this stalemate can be broken by empirical evidence either. We have good reason to think that non-interactionist property dualism and physicalism are equally compatible with all empirical evidence. For physicalism to be true, phenomenal properties must have “neural correlates”. This is the term that became widely used for the physical or functional basis of phenomenal experience in general and of specific phenomenal experiences. The correlation in question is supposed to be (at least) nomological\(^{52}\): a “neural correlate” of some phenomenal property is a physical or functional property that is coextensive with a phenomenal property in all creatures both natural and artificial in (at least) all nomologically possible worlds. Because the non-interactionist property dualist – like the physicalist – believes that the physical is causally closed, she is equally committed to the existence of such neural correlates. If the physical is causally closed, the phenomenal nomologically supervenes on the physical.\(^{53}\) So both physicalism and non-interactive property dualism is committed to the existence of neural correlates of consciousness. Empirical evidence so far seems to bear this out. Of course, it might turn out that the physical is not causally closed and they are both wrong on this issue; but nothing so far points strongly in this direction.

For physicalism to be true, it also has to be the case that the puzzling features of consciousness - the conceivability of zombies, the explanatory gap, and all the other epistemic gaps – like everything else, are actually physically explicable. Here is where the constitutional account comes in as an explanation of how phenomenal concepts work. The essence of the phenomenal concept strategy is to explain the epistemic gaps in terms of the structure of phenomenal concepts in a way that is compatible with physicalism. Of course, the constitutional account – or some other version of the phenomenal concept strategy – has empirical consequences; and its verification is contingent on the findings of future psychology and neuro-science. But I believe that we have strong reason to think that some such account \textit{will} be shown compatible with the empirical evidence. Even dualists tend to subscribe to the constitutional account\(^{54}\) (involving not the neural states, like physicalists think, but the phenomenal correlates of neural states). So it seems plausible that the stalemate between physicalism and dualism cannot be broken by future empirical evidence.

\[^{52}\] If the phenomenal property and the correlated physical property are the same property then the correlation holds in all metaphysically possible worlds but if dualism is true the correlation may hold only in worlds in which there are laws linking the distinct phenomenal and physical properties. The term “correlation” is neutral between metaphysical and nomological connection.


\[^{54}\] E.g., Chalmers 2003.
VI Indeterminacy

In closing, I would like to consider what this all means for the prospect of solving the mind-body problem. Let’s suppose for the sake of argument that non-interactive property dualism and physicalism are both empirically equivalent and philosophically on a par. By philosophically on a par I mean that they can each be developed so as to account equally well (or badly) for our philosophical intuitions and that there is no external, that is, non-question-begging, philosophical principle that can settle the dispute between them.

So far, we have assumed that there is a matter of fact that the dualist and the physicalist disagree about. It certainly seems as clear-cut a difference between the two accounts as they come: according to the physicalist, the fundamental properties and entities are exhausted by (non-mental) physical properties and entities while the dualist holds that there are some fundamental mental properties as well. It seems that there is crucial factual difference between the two accounts that is easy to grasp albeit one that cannot be empirically verified given that the mental and the physical are (at least nomologically) inseparable on both accounts. One might think that positing non-physical properties that are nomologically inseparable from physical ones, and that are epiphenomenal or always work in tandem with physical ones is so much excess metaphysical baggage. Nevertheless, it seems clear that there are two distinct theories in play.

This is a realist approach. It might be that the truth is beyond our reach about the mind-body problem; but that doesn’t necessarily call in question the facticity of the matter. There are some other instances of statements which seem to be either true or false but where seem to be neither a priori nor a posteriori ways of deciding the issue. For example, mathematical realists typically think that Cantor’s Continuum Hypothesis is an instance. But the case at hand seems different since it doesn’t involve the complexity of the mathematical case. A better analogy is the case of underdetermination of theory by evidence of which there are examples in the sciences.

But here is an alternative approach: perhaps physicalism and dualism are not really in disagreement. They just use different conceptual frameworks to get at the same facts. What I suggest is that the relevant notions of concept and property that figure in explicating these

---

55 From now on when I say “dualism” I mean non-interactive property dualism.

56 A case in point is the Bohmian versus the Everett versions of quantum-mechanics.
different metaphysical positions are indeterminate in a way that prevents the two positions to express a real difference. In other words, dualism and physicalism might encode a difference in the very conception of properties and their relationship to concepts, preventing the two positions to express a genuine factual difference. This suggestion might strike one as very counterintuitive; of all metaphysical disputes, the difference between physicalism and dualism seems as real as any. Yet indeterminism might be the most satisfactory solution of the metaphysical gridlock discussed in this paper.

What are the relevant notions of property in play?

a) *Dualism*. Dualists uphold the CP Principle. A principle about concepts and properties follows from a general version of the CP Principle:

1. If $C_1$ and $C_2$ are distinct concepts and both refer rigidly to the same property $P$ then $F & \sim C_1 = C_2$ (where $F$ is the complete fundamental truth) is inconceivable.

The version of the CP Principle that figures in the Zombie Argument states that if $P & \sim Q$ is conceivable then $P & \sim Q$ is metaphysically possible (where $Q$ is a phenomenal statement); but it is meant to hold with regard to any fundamental truth $F$ and any statement $T$ in a world, including of course identity statements. Since $C_1 = C_2$, if true, is necessarily true, it is not possible that it is false. But if $F & \sim C_1 = C_2$ is conceivable, then, according to the CP Principle, it is possible, so $\sim C_1 = C_2$ is possible. It follows then from the CP Principle that for any true statement $C_1 = C_2$, $F & \sim C_1 = C_2$ is inconceivable. This rules out, for example, phenomenal-physical identities – which, as the zombie case shows, can be conceived to be false consistent with the full physical truth – and so rules out illuminati as well, just as with the CP Principle.

Furthermore, the dualist conception of properties plausibly allows for properties that are not causally efficacious, to make room, given the causal closure of physics, for epiphenomenalism (and so as to avoid the implausible doctrine of causal overdetermination):

2 Properties need not be causally efficacious.

b) *Physicalism*. Physicalists reject the CP Principle and hold that

1’ Distinct concepts $C_1$ and $C_2$ may refer to a single property $P$ even though $F & \sim C_1 = C_2$ is conceivable.
2’ Properties are causally efficacious.

Our ordinary notion of property is pulled in different directions between the physicalist and dualist conceptions: it seems committed to both 1 and 2’. But given the completeness of physics, 1 and 2’ leads to the conclusion that there are no phenomenal properties, or that phenomenal properties and physical properties causally overdetermine their effects, both of which goes against common sense. The ordinary notion is then defective. In absence of a satisfactory unified concept of property, we oscillate between the dualist and physicalist conceptions: D-properties satisfy 1 and 2, and P-properties satisfy 1’ and 2’. If C₁ is a phenomenal concept it denotes a P-property that is also denoted by a physical concept C₂ and it denotes a D-property which is not denoted by any physical concept. Physicalism is true for P-properties and dualism is true for D-properties!

These conceptions figure in the most fundamental metaphysical accounts of reality and as far as I can see there are no independent standards to assess their respective merits. I suggest the reason for this might be that our concept of property is indeterminate between these two conceptions and there is no bit of reality that one latches on to that the other misses. On this view, the dispute between dualism and physicalism is not a substantive disagreement; it is a terminological disagreement. But it is a terminological disagreement of a peculiar sort.

It is a terminological dispute, for example, whether today is twice as hot as yesterday where the person using the Celsius scale agrees whereas the person using Fahrenheit doesn’t; or whether Trump is bald or hairy. But there is a difference between these cases and the mind-body case: in all of these cases there is a bedrock of more basic facts (number of hairs, etc.) that both sides agree on. They can be described at a more fundamental level in a way that can garner agreement from both sides of the Fahrenheit/Celsius, or bold/hairy divide. Even deep terminological issues like the issue of whether composite objects exist don’t seem to cut as deep as the mind-body dispute; one might redescribe the manner of their existence (as conventional, or non-fundamental, or whatever) in a way both sides can agree on despite their different use of the concept of existence. The mind-body case, by contrast, is resistant to such attempts. It is not amenable to common ground; attempts at reconciliation would seem preposterous. Unlike in these other cases, in the mind-body case there is no more basic conceptual scheme that can pick out the “metaphysical bedrock” against which the disagreement about properties can be resolved. It is a dispute about what concepts to use to describe reality at the most fundamental level. My proposal is that there is no real disagreement in fact; both schemes get at the same reality, in different ways.
Conclusion

Does my suggestion imply that every philosophical dispute that has come to a gridlock – and one might describe many disputes in this way – is, at bottom, a mere terminological dispute? Not necessarily. There are some that resemble the pattern I have described. The question, for example, whether tropes or universals are fundamental might be one. But there are other recalcitrant philosophical disagreements that are relevantly different. The debate over the nature of free will, for example, does involve different notions of freedom. But the parties to the debate can in principle agree that the other has a coherent notion of freedom – and only disagree about the importance and occurrence of each kind of freedom; about which one of them very well might be right – and the other wrong.
Appendix

Here is how one can deduce the inconceivability of illuminati from the Zombie Argument. I will now construct the Zombie \textsubscript{a priori} Argument whose conclusion is that phenomenal experience in a purely physical world is inconceivable. I argue that if the premises of the Zombie Argument are true then the premises of the Zombie \textsubscript{a priori} Argument are true. Along the way, I make some of the implicit assumptions of the Zombie Argument explicit.

The Zombie \textsubscript{a priori} Argument

The first premise is derived from premise 1 of the Zombie Argument:

1\textsubscript{a priori} It is a priori that \( \forall P \forall Q (P \& \sim Q \text{ is conceivable}). \)

1\textsubscript{a priori} involves a generalization over 1; this generalization is – as we will see shortly – necessary to construct the argument for the inconceivability of physicalism we are aiming for. \( P \& \sim Q \) is conceivable for \textit{any physical statement} – and not just the full physical description of our world – and for the same reason, having to do with the conceptual independence of phenomenal and physical descriptions in general. Because of the connection between a priority and conceivability, this means that for any \( P \), and any \( Q \), it is a priori that \( P \& \sim Q \) is conceivable. However, though it is a further claim, I think it is also knowable a priori that for any \( P \), and any \( Q \) \( P \& \sim Q \) is conceivable, and this is because the general reason for the truth of individual conceivability claims – the conceptual independence of phenomenal and physical descriptions – is knowable a priori.

The next premise is a generalization of premise 2:

2\textsubscript{a priori} It is a priori that \( \forall P \forall Q \) (if \( P \& \sim Q \) is conceivable then \( P \& \sim Q \) is metaphysically possible). (CP a priori)

2\textsubscript{a priori} is an extension of premise 2 for all physical statements \( P \), making explicit that premise

\textsuperscript{57} \( \forall \) is a substitutional quantifier, \( P \) is a statement variable for (logically consistent) physical statements, and \( Q \) is a statement variable for phenomenal statements.

\textsuperscript{58} The general formulation – \( \forall P \forall Q \) (if \( P \& \sim Q \) is conceivable then \( P \& \sim Q \) is metaphysically possible or Russellian monism is true) – follows from Chalmers’ Master Principle, i.e., that conceivability implies primary possibility, and the a priori truth that the primary and secondary proposition of phenomenal statements coincide, in the same way as premise 2 – which is a particular application of it
2, if true, is true a priori.

Premise 3 claims that if P & ~Q is metaphysically possible then physicalism is false. This cannot be known a priori for the simple reason that the premise is only - non-trivially - true if Q is true and P is the full, true physical description of the world - which is not knowable a priori. For the purposes of the a priori version of the argument I need to make these assumptions explicit.

3 apriori) It is a priori that if Q is true and \( \forall P \forall Q(P \& \sim Q \text{ is metaphysically possible}) \) then physicalism is false.

3 apriori follows from an a priori consequence of Phys, i.e., that if physicalism is true then \( \exists P \forall T \square(P \Rightarrow T) \).

From 1 apriori and 2 apriori it follows - by plausible principles of the logic of a priority - that it is a priori that \( \forall P \forall Q (P \& \sim Q \text{ is metaphysically possible}) \), and that, together with 3 apriori implies:

ILLUMINATI-INCONCEIVABLE: It is a priori that if Q is true then physicalism is false.

Here is the argument again:

The Zombie a priori Argument

1 apriori) It is a priori that \( \forall P \forall Q (P \& \sim Q \text{ is conceivable}) \).

2 apriori) It is a priori that \( \forall P \forall Q (\text{if } P \& \sim Q \text{ is conceivable then } P \& \sim Q \text{ is metaphysically possible}) \). (CP a priori)

3 apriori) It is a priori that if Q is true and \( \forall P \forall Q (P \& \sim Q \text{ is metaphysically possible}) \) then physicalism is false.

ILLUMINATI-INCONCEIVABLE: It is a priori that if Q is true then physicalism is false.

---

59 \( \forall \) and \( \exists \) is a substitutional quantifier, \( P \) is a statement variable for physical statements, \( T \) is a statement variable for true positive statements, and \( \square \) is the metaphysical necessity operator.
References:


Papineau, D., 2002. Thinking about consciousness, Oxford UP.


