

## The Rise and Fall of the Mind-Body Problem<sup>1</sup>

Katalin Balog

**Abstract:** In this paper, I examine the relationship between physicalism and property dualism in the light of the dialectic between anti-physicalist arguments and physicalist responses. Upon rehearsing the moves of each side, it is hard not to notice 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, and it seems that there are neither a priori nor a posteriori grounds to choose between the two. I suggest that the reason for the intractability of the disagreement, perhaps surprisingly, is they are both true: physicalism and dualism are formulated in terms of different conceptual schemes, each involving basic metaphysical concepts such as property, law, possibility, and necessity. My proposal is that this means that there is no real disagreement in fact; both schemes get at the same reality, in different ways.

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 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, physicalists have reasonable, though not decisive rebuttals; they also have some arguments of their own against dualism.<sup>2</sup> A dispassionate observer might find an impasse between physicalism and anti-physicalism that is hard to break.

---

<sup>1</sup> I would like to thank Ned Block, David Chalmers, Troy Cross, Jenann Ismael, Barry Loewer, Raymond Martin, Howard Robinson, Gilad Tanay, Tobias Wilsch, participants of a discussion group at NYU, an NYU workshop on *Grounding and Consciousness*, audiences at the Rutgers University philosophy department, students at Jenann Ismael's philosophy of mind seminar, and audiences at conferences at CEU (CEU-Rutgers Mind Workshop), NYU (NYU Institute for Mind, Brain, and Consciousness), and at the *Mind Brazil International Workshop* in Tiradentes, Brazil for comments and helpful discussion of this paper.

<sup>2</sup> The main arguments in favor of physicalism 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.

In the first part of the paper, I discuss the anti-physicalist arguments and physicalist replies to them. The anti-physicalist arguments start from a premise about a conceptual, epistemic, or explanatory *gap* between physical and phenomenal descriptions<sup>3</sup> and conclude from this – on *a priori* grounds – that if the phenomenal is real physicalism is false.<sup>4</sup> Of course these arguments support anti-physicalism only on the assumption that phenomenal experience really does exist – which is increasingly being challenged by illusionism<sup>5</sup> – but I want to put illusionism to the side in this paper. My first aim is to develop a master argument that rebuts the anti-physicalist arguments. The master argument is based on what has come to be known as the phenomenal concept strategy<sup>6</sup>; this strategy – following Brian Loar’s original proposal in (1990/1997) – appeals to the special cognitive features of phenomenal concepts 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 defend itself from attacks from the other position and view the other position as question-begging. There are neither *a priori* ways to decide between the two, nor are there *a posteriori* ways – pending shocking scientific discoveries about the independent causal powers of the mind –, rendering the question

---

<sup>3</sup> 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 Dennett (2016), and Frankish (2016), as well as Rey (1996), Graziano (2013), Humphrey (2011), and Perebrom (2011), though the latter are more equivocal about their position.

<sup>4</sup> These arguments include, among others, arguments based on conceivability considerations by Kripke (1972), Nagel (1974), Bealer (1994), Chalmers (1996, and 2009), as well as the Knowledge Argument of Jackson (1982), versions of the Property dualism Argument in Robinson (1993), White (2007), and Nida-Rümelin (2007), and the Explanatory Gap Argument in Levine (2001) and (2007).

<sup>5</sup> See, e.g., Frankish (2016).

<sup>6</sup> Stoljar’s (2005) phrase.

unanswerable.

I conclude the paper by proposing a novel way to look at this impasse: maybe the question of whether consciousness is purely physical or not does not track real alternatives. Maybe physicalism and dualism are both true; they describe the same reality in radically different vocabularies. This, by the way, means that I take other positions on the mind-body problem: substance dualism, interactive property dualism, idealism, illusionism, and various forms of neutral monism to be false; though I will not argue in detail for this in this paper.

## **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.<sup>7</sup> Intentionality and consciousness is instantiated in macroscopic systems *in virtue of* immensely complex arrangements of fundamental physical properties and entities, i.e. for biological individuals in virtue of brain states and 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

---

<sup>7</sup> 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.

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 truths in that world, including, as the case may be, truths about phenomenal consciousness,<sup>8</sup> are metaphysically necessitated by the complete physical truth about that world.<sup>9</sup> Accordingly, if physicalism is true in our world, the *Physicalist Entailment Thesis* is true as well:

(Phys)  $\forall T \Box (P \supset T)$ .<sup>10</sup>

It follows that if there are truths about phenomenal consciousness – for example, that Mary knows what it is like to see red – that are not necessitated by the complete physical description  $P$  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, neutral monism,

---

<sup>8</sup> Jackson stipulates that it is only *positive* truths that have to be necessitated by the complete physical truth. 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 than 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.

<sup>9</sup> 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 might not express the full physicalist commitment – only a necessary condition – because it is apparently compatible with a non-physicalist ontology according to which there are fundamental mental as well as fundamental physical properties connected by “brute” necessary connections.

<sup>10</sup>  $\forall$  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

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.<sup>11</sup> 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. This leaves non-interactive property dualism and panpsychism as the strongest candidates.

Non-interactive property dualism is simply dualism committed to the causal closure of physics – not necessarily a view denying mental causation. It holds that the physical does not necessitate all truth, and that the fundamental ontology of the world includes phenomenal properties. It also posits the existence of fundamental vertical laws<sup>12</sup> that connect arrangements of physical entities and properties to phenomenal properties. Consequently, it is committed either to epiphenomenalism or causal overdetermination of the physical by the mental. Since I find epiphenomenalism too obviously implausible, I will take non-interactive property dualism as being committed to causal overdetermination.

Panpsychism is back in vogue due to the frustration over the impasse between physicalism and

---

<sup>11</sup> However, see, e.g., Nida-Rümelin 2017 for an argument for mental substance.

<sup>12</sup> In principle there could be a non-naturalistic non-interactive property dualism that denies the existence of vertical laws between the physical and the phenomenal, but it has not been advocated; probably due to the insuperable difficulties of explaining why behavior and consciousness remain correlated in humans. 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.

property dualism.<sup>13</sup> Its most plausible version, Russellian neutral monism, holds that the intrinsic, as opposed to causal/dispositional, nature of fundamental properties and entities is phenomenal. 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) and Strawson (2006), e.g., considers neutral monism a version of physicalism. Nothing important hangs on the terminology; however, I think it is useful to firmly distinguish Russellian neutral monism from the version of physicalism I will be mostly concerned with in this paper, according to which physicalism requires that fundamental physical properties and entities are non-mental. On this definition, physicalism and Russellian panpsychism are incompatible.

Russellian monism, too, is committed to the causal closure of physics since it holds that the causal dispositional nature of fundamental properties is physical. It claims to solve the problem of mental causation that plagues non-interactive property dualism; according to it, the intrinsic nature of the causally relevant physical properties involved is itself mental. Papineau (2020), however, argues that it ultimately does no better than property dualism explaining mental causation. The central consideration cited in favor of Russellian monism is the claim that physics as it is is incomplete; it doesn't tell us what physical things are like *in themselves*, intrinsically, so to speak. This is a highly controversial claim. 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 at least as hard to see how the experiences of micro-entities in the brain could combine to give rise to a composite experience had by a human being as figuring out how micro-physical property instantiations in the brain give rise to the same experience.

---

<sup>13</sup> See, e.g., Chalmers 2017, Goff 2017, Mørch 2019, Strawson 2006.

Russellian monism is too speculative and implausible, without improving on physicalism with respect to the explanatory gap. Similarly, I rule out substance dualism as we do not have good reasons for believing in a mental substance.<sup>14</sup> Putting idealism to the side, that leaves physicalism and non-interactive property dualism<sup>15</sup> with causal overdetermination as the best candidates. These two views seem to provide starkly different ontological pictures of the world. The remit of this paper is to convince you that this might not be the case after all.

## II. The zombie argument and the inconceivability of purely physical minds

There is a line of argument against physicalism that goes back at least to Descartes' argument for the distinctness of mind and body. These arguments conclude, based on *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"; which is based on the idea that zombies, i.e., creatures that are physically identical to normal humans but have no phenomenal experiences whatsoever, are conceivable. Here is a more general formulation of the argument.

### *The zombie argument*

- 1)  $P \leftrightarrow \sim Q$  is conceivable.<sup>16</sup>
- 2) If  $P \leftrightarrow \sim Q$  is conceivable then  $P \& \sim Q$  is metaphysically possible (CP principle)
- 3) If  $P \& \sim Q$  is metaphysically possible then physicalism is false.

---

<sup>14</sup> See Balog 2022.

<sup>15</sup> From now on, unless I otherwise specify, by "dualism" I will mean non-interactive property dualism with causal overdetermination.

<sup>16</sup>  $P$  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.

---

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”<sup>17</sup>. A statement is *positively conceivable* if it is possible to form some sort of clear and distinct conception of a situation in which the statement is true. In Balog (1999), I have provided a reductio for the original version of Chalmers’ zombie argument in which he appeals to the *negative conceivability* of zombies (a statement is negatively conceivable if it is not possible to rule it out on a priori grounds). Running the zombie argument based on the notion of positive conceivability renders it immune to the zombie refutation in Balog (1999), given that it is based on the negative, and not the positive conceivability of certain states of affair for zombies. Even though it is not all that clear what having a clear and distinct conception amounts to, the physicalist is well advised to come up with another counterargument.

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 consciousness reject premise 1.<sup>18</sup> But I agree with Chalmers that no physical description *a priori* entails any 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

---

<sup>17</sup> Chalmers introduces a battery of different conceivability concepts in (2002a), among them negative conceivability, positive conceivability, ideal conceivability, etc.

<sup>18</sup> E.g. Lewis 1966 and Jackson 2003, 2007, 2020. Analytic functionalism or representationalism concerns the *meaning* of phenomenal terms; it says that such meanings can be analyzed in functional or representational terms. *Pain*, e.g., according the analytic functionalism, has a conceptual role that connects it (in the meaning-constituting way) with complex concepts like *typically caused by injury*, *typically causes avoidance behavior*, *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.



the crucial premise in the argument is 2.<sup>19</sup>

In fact, the anti-physicalist arguments are all based on a premise concerning the epistemic status of the phenomenal with respect to the physical (e.g., that zombies are *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 connote 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. Is the conclusion of the zombie argument a priori knowable then? One might object that one couldn't know premise 3 without knowing – a posteriori – that Q is true. However, we can rewrite the argument in the following form:

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

Since all three premises can be known a priori (if at all), the conclusion can also be known a priori.

But 4 can be rewritten as

---

<sup>19</sup> 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 discuss what I think goes wrong with a related argument by Frank Jackson (Balog 2001).

5) If Physicalism is true then Q is false.

In other words, the zombie argument, if sound, rules out, on a priori grounds, the existence of purely physical minds – that is, purely physical duplicates of normal humans who are at the same time conscious.<sup>20</sup>

My strategy to counter the conceivability arguments is to argue, to the contrary, that the existence of purely physical minds cannot be ruled out a priori, that they are conceivable, at least in the negative sense. This is enough to counter the zombie argument.<sup>22</sup>

This will not amount to a reductio of the zombie argument; my aim is simply to argue, on independent grounds, for a claim that is not compatible with its conclusion. Given that the argument is valid this means that one of its premises must be false. I will also argue that the evidence points at premise 2 as the culprit.

### III. Purely physical minds

Showing that purely physical minds are conceivable works not only against the zombie argument, it works against the other anti-physicalist 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 conclusion that if phenomenal experience exists then physicalism is false.

But how can the physicalist argue for the negative conceivability of purely physical minds? It will be instructive to compare what is involved in the negative conceivability of zombies as opposed to the

---

<sup>20</sup> Note that “purely physical” is not meant to exclude the presence of mental properties; only the presence of *fundamental* mental properties; so anti-zombies are not *trivially* inconceivable.

<sup>21</sup> 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.

<sup>22</sup> Physicalists, of course, are committed to the *actual existence* of purely physical minds. But the argument will only require their conceivability.

negative conceivability of purely physical minds. 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<sup>23</sup> yet direct way, i.e., there are no analytically sufficient conditions in physical, functional, or behavioral terms for their application; this results in enough conceptual independence of phenomenal and physical concepts to make zombies negatively conceivable. How about purely physical minds? The basic idea is this. Zombies are negatively conceivable because we grasp phenomenal properties in a substantial yet direct way<sup>24</sup>; but purely physical minds 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. In a physicalist version of panpsychism, one might even claim that stars or rocks are conceivably conscious in virtue of their physical properties. You can object 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 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 such property identities is not thereby conceiving of physicalism being true. Perhaps there are possible worlds where physicalism is false for reasons having nothing

---

<sup>23</sup> 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.

<sup>24</sup> 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.

to do with phenomenal properties (because of the instantiation of some non-phenomenal non-physical property). However, I don't see any reason why, in conceiving of these property identities, one couldn't just also conceive that there are no further non-physical, or non-physically realized properties.

So far so good. But there is a problem. Even if – as I argued – purely physical minds cannot be ruled out on conceptual grounds, perhaps they can be ruled out via the CP principle (Premise 2 of the zombie arguments)? It seems question-begging for the physicalist to rest their case about the conceivability of purely physical minds 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 in defense of physicalism.<sup>25</sup> One way to do that would be to show that purely physical minds are not only negatively conceivable, but also *positively* conceivable – which is the relevant notion appealed to in the latest iteration of the zombie argument. There are several authors who offered arguments along these lines. Frankish (2007) introduces the notion of anti-zombies – purely physical conscious beings – to construct a reductio: he starts with the premise that anti-zombies are conceivable, and concludes that they are possible, given the CP principle. The strategy is to add plausible premises to the zombie argument to derive contradictory conclusions, thereby showing some of its premises to be false. Sturgeon (2000), Brown (2010) and Piccinini (2015) similarly argue that conceivability cuts both ways and so that the zombie argument undermines itself. Though they do not discuss the distinction between positive and negative conceivability, I think their claims should be interpreted as involving positive conceivability, considering that Chalmers' latest and best version of it employs positive conceivability as well. In effect, there is some reason to think that the positive conceivability of

---

<sup>25</sup> 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. As I mentioned before, this argument doesn't work against the most recent formulations of the zombie argument in terms of positive conceivability.

zombies and the positive conceivability of anti-zombies 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 anti-zombies (in the way she might try to use it against the negative conceivability of anti-zombies): if the dualist argues that the positive conceivability of purely physical minds can be ruled out a priori on the basis of the positive conceivability of zombies *together with* 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 purely physical minds *together with* the CP principle to rule out the positive conceivability of zombies.<sup>26</sup>

However, as I mentioned, there are questions about the positive conceivability of purely physical minds – that is, whether we have a clear and distinct conception of them – 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* conceivability of purely physical minds help? So far what we have seen is that the physicalist can show that purely physical minds 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. They might, however, be able to *bolster* their argument for the conceivability of purely physical minds by not only considering the conceptual coherence of psycho-physical identities but considering it in the context of an account of properties, modality, 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. Such a physicalist account would at the same time explain what goes wrong with the CP principle from which the anti-physicalist arguments arise, and so would neutralize its threat against the

---

<sup>26</sup> An argument with a similar structure is suggested in Loewer (1978) with respect to Leibniz's Ontological Argument.

conceivability of purely physical minds.<sup>27</sup> I will now turn to this account.

#### IV. The Phenomenal Concept Strategy

As the crucial ingredient of such an account, 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*

It turns out we can come up with a story of how introspection works – how we can become aware of our phenomenal states in a direct and substantial way – that is fully compatible with physicalism. We can explain how, even if phenomenal states are purely physical (or physically realized), there could be a physical mechanism that accounted for the introspection of these states and that would reliably produce all the epistemic puzzlement we find ourselves with regarding consciousness. Such an account would support the conceivability of purely physical minds by explaining how the CP principle can be false.<sup>28</sup> My favored version of the phenomenal concept strategy is the constitutional account of phenomenal concepts. According to it,<sup>29</sup> there is an intimate relationship between phenomenal concepts and their referents; token experiences serve as modes of presentation of the

---

<sup>27</sup> Hill and McLaughlin (1999) makes a similar proposal in the context of an argument for physicalism.

<sup>28</sup> See, e.g., recognitional accounts (Tye 2003), demonstrative accounts (Levine 2007, Perry 2001), and information-theoretical accounts (Aydede & Güzeldere 2005).

<sup>29</sup> I defend a version of that account in Balog 2012a. Other versions of the constitutional account have been proposed by physicalists such as Hill and McLaughlin (1999), Block (2007), and Papineau (2002, 2007); and dualists like David Chalmers (2003).

phenomenal properties they instantiate.<sup>30</sup> In the case of most concepts, e.g., the concept WATER, it doesn't matter whether the neural configurations that constitute a particular token of WATER are partly constituted by water or not. But in the case of phenomenal concepts, e.g., the concept PAIN, constitution matters for reference, both in terms of how the concept cognitively presents its reference, and in terms of how the reference is determined. More precisely, on this view, every token of a phenomenal concept applied to current experience is constituted by *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 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, *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.<sup>31</sup> Of course, for the constitutional account to work, some of its details have to be worked out, in particular, a physicalistically respectable story has to be told of how constitution can play a role in determining reference.<sup>32</sup>

The constitutional account of phenomenal concepts explains how the epistemic/conceptual gaps between the phenomenal and the physical are compatible with physicalism by providing a physicalist model of how we can have a direct, yet substantial grasp of phenomenal properties. First, the constitutional account explains how we can have a *substantial grasp* of phenomenal properties even while this grasp is direct, and unmediated by physical or functional modes of presentation. Because

---

<sup>30</sup> In what follows I concentrate on phenomenal concepts that refer to phenomenal *properties*, but the account can be easily modified to apply to concepts that refer to *particular instances* of phenomenal states.

<sup>31</sup> 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).

<sup>32</sup> 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). Also, I am disregarding here the general problem of a physicalist account of reference.

in the canonical, first person, present tense applications of a phenomenal concept an instance of the property it refers to is *literally* (physically) present in the concept, there will be always something it is like to token the concept in those canonical applications.<sup>33</sup> Every instantiation of a phenomenal property reveals something essential about that property, namely, it reveals *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 tokens of phenomenal concepts present their referent as the property whose token they incorporate – and not via any functional or physical description – they will refer to phenomenal properties *directly*, as well as substantively.

Nothing in this account is incompatible with physicalism. Yet on this account it follows that, because of the directness and substantiality of our direct phenomenal concepts, zombies are conceivable.<sup>34</sup> Assuming that physicalism is true, we can now explain the conceivability of zombies as well as the failure of the CP principle.

The CP principle might hold in all sorts of cases, it might even hold in all cases except the phenomenal case. Phenomenal truths are supposed to be different from truths involving concepts like WATER, or BRAIN STATE B. Chalmers and Jackson (2001) claim that these concepts are associated *a priori* with descriptions (e.g. “the transparent potable liquid...”, “oscillations in the V4 area of visual cortex”) and they also claim that these connections are *sufficient* to obtain *a priori* entailments from the full physical description to all statements of fact, e.g., they are sufficient to rule out a priori a scenario where everything is physically the same but yet there is no water or brain state B. A physicalist might go along with this much, and affirm that for all truths T but phenomenal ones,

if  $P \& \sim T$  is conceivable then  $P \& \sim T$  is metaphysically possible.

---

<sup>33</sup> Levine (2006, 2007) is critical of this approach. He argues that it is impossible to explain *cognitive presence* by *physical presence*.

<sup>34</sup> For a detailed account of how this works see Balog 2012a.



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.<sup>35</sup> I am going along with the CP principle – except for phenomenal statements – for the sake of argument; the CP principle can be rebutted in a way that doesn't depend on rejecting its semantic framework wholesale.<sup>36</sup>

The point is that even if one accepts the CP principle with regard to all true non-phenomenal statements, one might very well reject it with regard to phenomenal statements. Because of the fundamentally different cognitive architecture of phenomenal concepts, there are no *a priori* connections between phenomenal and physical/functional/structural concepts that are sufficient to rule out the zombie scenario *a priori* – 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 individuating of them. I think it is quite plausible that there are conceptual links, even perhaps concept individuating 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, as well as perhaps concepts related to cognitive architecture, as I pointed out before. 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. The point remains that because phenomenal concepts are direct and substantial, we can conceive of zombies. This explanation is equally compatible with a physicalist or a dualist metaphysics and leaves

---

<sup>35</sup> 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 *a priori entailment thesis* and Chalmers's Master Principle that lies behind the CP principle, see my doctoral thesis (Balog, 1998), p 124-5.

<sup>36</sup> But see, e.g., Block and Stalnaker (1999), McLaughlin (2007) and Schaffer (2017) for an argument that these entailments are not *a priori* even for positive non-phenomenal statements.

the CP principle in place for all except phenomenal statements.

Although I won't go into the details, it is easy to see that 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, or the semantic stability of phenomenal concepts.<sup>37</sup>

## **V. Metaphysical gridlock?**

Let's consider where this leaves the dialectic between the physicalists and dualists. The phenomenal concept strategy notwithstanding, it must be granted that the dualist can still consistently claim that, as opposed to zombies, purely physical minds are not conceivable given the CP principle. I will get into more detail about the alternative metaphysical frameworks in the next section. For now, let's briefly recap the main arguments in favor and against these two positions.

The main argument in favor of physicalism concerns mental causation and the causal closure of physics. Loewer (1995) and Papineau (1995) 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. Chalmers, on the other hand, argues (2002a, 2002b) that the CP principle – which he uses to argue for dualism – provides the simplest and most satisfying account of modality.

Neither of these considerations seem to definitively outweigh the other. In particular, both the physicalist and the dualist can answer objections from the other side. By their own lights, the physicalist has rebutted the anti-physicalist arguments. From their point of view, the CP principle is

---

<sup>37</sup> For a more detailed account of all of these, see Balog (2012a).

shown to be wrong while its intuitive appeal can be admitted. A physicalist account of modality can be furnished. A simple and straightforward account of mental causation is on offer. However, the dualist is in a similar situation. According to the dualist, for all the clever verbal feats of the phenomenal concept strategy, purely physical minds are still inconceivable, as shown by the CP principle. The CP principle also underlies a satisfyingly simple account of modality. Causal overdetermination, while not intuitive, provides an acceptable account of mental causation. Let's pause here for a moment.

On the one hand, the dualist 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 purely physical minds are conceivable – based on the phenomenal concept strategy – and uses this claim to undermine the CP principle. By their own lights both sides seem justified in holding on to their key principles and denying those of their opponents. Where you end up depends on what you take as your starting point.

This is puzzling. One would have thought that when it comes to a priorities like the CP principle, or the conceivability of purely physical minds, there are a priori ways to justify or refute them, in a way that is independent of one's theoretical commitments. But it seems like there are no neutral principles *outside* the physicalist and dualist systems that could settle this issue.

The situation would be different if there was the possibility of empirical disagreement. But given that both views are committed to the causal closure of physics, such disagreement is out of the question.<sup>38</sup> All action has, on both views, a fully physical explanation. As for the mind-brain connection, physicalists and dualists agree that phenomenal properties have “neural correlates”. The term “correlation” is neutral between metaphysical and nomological connection. The physicalist

---

<sup>38</sup> On the assumption that causal closure holds, that is. If it doesn't interactive dualism wins.

believes that phenomenal properties and their correlated physical properties are identical, or that the physical properties realize the phenomenal properties. On the other hand, assuming causal closure, the dualist is committed to vertical nomological relations between the physical and the phenomenal.<sup>39</sup> The physicalist and the dualist will expect the same correlations to hold, their only difference being whether they consider these correlations to be metaphysical or nomological in nature. As long as the causal closure of physics holds, there is no possible empirical evidence that would count in favor of one or the other of these views – all evidence that is compatible with the one is compatible with the other as well.

In the absence of decisive priori arguments, or empirical evidence for either side, one might think it comes down to the overall merits of the two systems of metaphysics, like their handling of the problem of mental causation and the like. But it is hard to see how one would go about weighing the pros and cons as they seem incommensurable.<sup>40</sup> Let us suppose for the sake of argument that dualism and physicalism are philosophically on a par. By philosophically on a par, I mean that they can each be developed so as to give a satisfactory overall metaphysics of the mind and the world and that there is no non-question-begging philosophical principle that can settle the dispute between them. I would like to consider a new way to understand the mind-body problem in the light of this.

## **VI Physicalism versus dualism: a terminological issue?**

All along, 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 disagreement over fact between the two accounts as they

---

<sup>39</sup> See Loewer (1995).

<sup>40</sup> But see Hill and McLaughlin (1999), and McLaughlin (2007) who argue 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.

come: according to the physicalist, there are no non-physical fundamental properties and entities while the dualist holds that the fundamental properties include mental properties as well. It looks like either dualism is true, and physicalism is false, or vice versa. The factual difference between the two accounts appears easy to grasp even if – as it appears to be case – there is no empirical or philosophical method to adjudicate between them. Though philosophers have traditionally been split between dualists and physicalists, arguing for the truth of their respective positions, I suspect most would remain convinced that there is a fact of the matter which of dualism and physicalism is true even if they believed this fact to be unknowable. The position that there is a truth to the matter even if it might be beyond our comprehension, has intuitive appeal.

There are other instances of statements which seem to be either true or false but where there seem to be neither a priori nor a posteriori ways of deciding. For example, mathematical realists typically think that Cantor's Continuum Hypothesis is an instance. But the case at hand seems different since it does not involve the complexity of the mathematical case. A better analogy is the case of underdetermination of theory by evidence of which there seem to be examples in the sciences. While physicalism and dualism are philosophical, rather than strictly speaking scientific theories, it will be instructive to look at how the problem appears in science.

There are two positions about the epistemic status of empirically equivalent scientific theories. One is the God's eye view, the position I have just mentioned with regard to the mind-body problem, according to which even though the truth might be beyond our reach, there are factual differences in how the two theories describe the world. This is how most philosophers would describe the difference between the Bohmian and the Everett interpretations of quantum-mechanics, even on the – controversial – assumption that they are empirically equivalent.<sup>41</sup> They have a different description

---

<sup>41</sup> The problem with the account is whether it can account for probabilities. If so, it is empirically equivalent with Bohm's theory, even though they differ in their ontology.

of what the fundamental entities, properties, and laws are. A more controversial example discussed in Van Fraassen 1980 involves variants of classical mechanics differing only that the one assumes the entire universe to be at rest and the other assumes instead that it is moving with some constant velocity in some direction. There is no way for us to detect constant, absolute motion or absolute rest by the universe as a whole, and so there exist all these infinitely many theories with the exact same empirical consequences. These theories appear to say different things about the universe, though it might be argued that they contain "surplus structure" and that the theory should be formulated in a way that preserves its empirical consequences while eliminating any such surplus.

There has been a lively debate among philosophers under what circumstances empirically equivalent theories that apparently say different things should be nevertheless considered to agree.<sup>42</sup> Quine's 1975 suggests the following criterion for theoretical equivalence:

...two formulations express the same theory if they are empirically equivalent and there is a reconstrual of predicates that transforms the one theory into a logical equivalent of the other. (320)

Quine's toy example is electron theory, and a theory we get when we switch the words 'molecule' and 'electron'. The two theories are empirically equivalent, and though contradictory, can be reconstrued in a way that makes them logically equivalent. Van Fraassen's variants of classical mechanics is a plausible example of theoretical equivalence under this criterion, as is the Newtonian and Hamiltonian versions of classical mechanics as their fundamental ontology can be cross-defined, and their respective laws describe the same evolution of that fundamental ontology. The Bohmian and the Everett formulation of quantum mechanics would come out as different theories under this

---

<sup>42</sup> See, e.g., Reichenbach 1961, 374-75, Putnam 1963, and Salmon 1966 for the view that empirically equivalent theories must always say the same thing about the world. For the opposite view see Glymour 1970.

criterion.

In the case of physicalism and dualism it seems clear that such a “reconstrual” is not possible. There is no suitable substitution of predicates that could turn the claim, e.g., that brain states and phenomenal states are connected by vertical law into the claim that they are connected by metaphysical necessity; there is no simple isomorphism between the two theories. This would support the intuitively plausible idea that physicalism and dualism do say different things about the world, despite their empirical equivalence.

What I will argue in the rest of the paper is that this assessment is mistaken. There is another way in which two theories can say the same thing even if they apparently contradict each other. The central concepts which dualism and physicalism uses to describe the world at its most abstract are different, and, as a result, they provide superficially contradictory descriptions of the world without really contradicting each other. The crucial notions that vary in dualism and physicalism are the notions PROPERTY and METAPHYSICAL NECESSITY. If one adopts one conceptual scheme, dualism comes out as true. If one adopts another conceptual scheme, physicalism comes out as true.

My proposal, which is akin in spirit to Putnam’s conceptual relativity,<sup>43</sup> is a kind of anti-realism that doesn’t altogether deny the idea of an objective world but asserts that the world doesn’t have an “innate” structure to be tracked by our conceptual schemes but rather, it is possible to be understood and described equally well in terms of alternative conceptual schemes. The idea is that dualism, which marks a difference that makes no difference – i.e., that phenomenal states are distinct from, but nomologically inseparable from brain states, and always work, causally speaking, in tandem with them – simply carves up the world in different ways than physicalism but there is no reason to suppose that one way of carving up the world really corresponds to how things are while the other

---

<sup>43</sup> See, e.g., Putnam 1982 for the general idea of no “ready-made world”.

doesn't. The metaphysical realist, the proponent of the God's eye view of the world will, of course, disagree. They will say that at most one of these conceptual schemes can correspond to how what properties and modality really are. While one can reasonably disagree about this, my position is not ruled out by anything we know about the world. It would also explain why the mind-body problem appears so exasperatingly intractable.

I don't mean that disagreements about the mind-body problem are merely terminological in the usual sense. 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 these cases, there is a bedrock of more basic facts (number of hairs, etc.) that both sides agree on. There is no sense of mystery about the disagreement. They can be described at a more fundamental level in a way that can garner agreement from both sides of the Fahrenheit/Celsius, or bald/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. 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 thesis doesn't imply is that there is anything indeterminate about reality, or the world.<sup>44</sup> Even if it turns out that quantum-mechanics says that physical reality is indeterminate in all sorts of ways, that would not be the issue here. I am not suggesting either that there is something *vague* about the

---

<sup>44</sup> For a discussion of metaphysical or ontic indeterminacy see Williams 2008.



concept PHYSICAL, or PROPERTY, or NECESSITY, etc. What I am saying is that both of these accounts describe the world, and they are both true, as far as we know, despite the fact that they employ different concepts of the basic building blocks of the world such as property, necessity, and law.<sup>45</sup>

Let's recap the two accounts. Whereas the physicalist believes that phenomenal properties are identical to, or realized by physical properties, the sort of dualist we are considering in this paper – committed to both causal closure and mental causation – believes that there are vertical nomological relations between the physical and the phenomenal.<sup>46</sup> This difference in views about fundamental ontology is a consequence of the dualist asserting, and the physicalist denying the CP principle, delineating the relationship between properties, concepts, and modality.<sup>47</sup> There have been other formulations of the principle. Nida-Rümelin 2007, e.g., claims that a person who grasps<sup>48</sup> the same property via different concepts can in principle find out without further empirical investigation that the two concepts co-refer. These are all versions of the transparency thesis linking properties, concepts, and modality first articulated in Descartes's *Meditations* (1641/1985) and formulated in one or another form by each anti-physicalist argument since. Dualism relies on a family of closely linked accounts.<sup>49</sup>

---

<sup>45</sup> A good candidate for metaphysical accounts that might be both true in this way is the dispositional/categorical debate about fundamental properties. A different suggestion with regard to the nature of fundamental properties is developed in Loewer 2020.

<sup>46</sup> See Loewer 1995.

<sup>47</sup> Chalmers 2009 adds some clarifications to the original zombie argument. He employs the two-dimensional semantic framework to characterize “primary possibility” and “secondary possibility” and argues for his Master Principle that conceivability implies primary possibility. The CP principle follows from this, on the plausible assumption that both P and Q express the same primary and secondary proposition. On the other hand, if we assume that P expresses different primary and secondary propositions and assume that P&~Q is not possible, Russellian monism follows. I will not follow out the ramifications of this for the two positions as I have already put Russellian monism aside.

<sup>48</sup> She uses the term to mean conceiving of a property via its essence.

<sup>49</sup> Another example of a different formulation to the same effect: no two concepts can share both their first and second intension. If C1 and C2 are distinct concepts, and both refer rigidly to the same property P then  $F \& \sim C1 = C2$  (where F is the complete fundamental truth) is inconceivable. All of these formulations offer a certain way of thinking about properties, concepts, and necessity.

What is the epistemic status of the CP principle and related other dualist principles? It is supposed to be an a priori truth, but of a peculiar sort. It is different from stock examples of a priori truths, e.g.,  $2+2=4$ , or BACHELORS ARE UNMARRIED, whose falsity can be ruled out by a priori methods independent of holistic considerations regarding one's overall conceptual framework. The a priori status of the CP principle, in contrast, seem to depend on holistic considerations. 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 conceivability arguments exploit. However, as I have argued above, there are no holistic considerations that weigh decisively in favor of the CP principle, and therefore it cannot demand assent from all parties to the debate – like  $2+2=4$ , or BACHELORS ARE UNMARRIED can – independent of their overall outlook.<sup>50</sup> I propose instead that the CP principle is a conceptual truth that articulates a connection between a network of bedrock concepts such as PROPERTY, CONCEPT and NECESSITY in a particular way. The dualist notion of PROPERTY, in particular, is committed to the idea that whenever there are two distinct concepts that share their first and second intensions they cannot corefer. In other words, we have special insights into the nature of properties by virtue of forming special concepts of them.

The physicalist, by contrast, has different concepts of PROPERTY, CONCEPT and NECESSITY. For the physicalist, there is no problem in the assumption that we could have two psychologically different concepts (say, a direct phenomenal concept, and a neurophysiological concept) each with coinciding first and second intensions, grasping the very same referent in different ways. The phenomenal concept strategy provides a model of how this can happen. Appealing to an alternative proposal of how concepts, properties, and necessity are related, the physicalist can account for all of reality while denying the Fregean assumption behind the transparency theses that there is a single entity that both individuates concepts and fixes their reference. This is a conception of properties that does not presuppose anything about insights into the nature of properties. It rather construes properties as

---

<sup>50</sup> I am indebted to Troy Cross for discussion on this issue.

the kind of thing that enters the causal order of things (and as a result, the physicalist can explain mental causation better than the dualist can).

There is a standoff between the two accounts. If what I suggest is right, the disagreement between the physicalist and the dualist<sup>51</sup> is akin to a terminological disagreement. But it is a terminological disagreement of a peculiarly deep sort that leads to a sense of mystery.

### References:

- Aydede, M. & Güzeldere, G., 2005. "Cognitive architecture, concepts, and introspection: An information-theoretic solution to the problem of phenomenal consciousness", *Noûs*, 39 (2): 197-255.
- Balog, K., 1998. *Conceivability and Consciousness*. Dissertation, Rutgers University.
- Balog, K., 1999. "Conceivability, Possibility, and the Mind-Body Problem", *The Philosophical Review*, Vol 108, No 4 (October 1999). Reprinted in Volume XXIII of *The Philosopher's Annual*.
- Balog, K, 2001. "The A Priori Entailment Thesis", Commentary on Frank Jackson's *From Metaphysics to Ethics*, *Philosophy and Phenomenological Research*, Vol. LXII, No. 3.
- Balog, K, 2009. "Phenomenal Concepts". In Brian McLaughlin, Ansgar Beckermann, and Sven Walter (Eds.), *The Oxford Handbook of Philosophy of Mind* (pp. 292-312). Oxford: Oxford UP.
- Balog, K, 2012a. "Acquaintance and the Mind-Body Problem". In Christopher Hill and Simone Gozzano (Eds.), *The Mental, the Physical* (pp. 16-43). Cambridge: Cambridge University Press.
- Balog, K, 2012b. "In Defense of the Phenomenal Concept Strategy". *Philosophy and Phenomenological Research*, 84(1), 1-23.
- Balog, K, 2022. The illusion of the enduring self. In M Nida-Rümelin, J Bugnon (eds.), *The Phenomenology of Self-Awareness and the Nature of Conscious Subjects*, Routledge.

---

<sup>51</sup> Bear in mind, in this paper we are only concerned with non-interactive over-determinist property dualism.

- Bealer, G. (1994). "Mental Properties," *Journal of Philosophy* 91:185-208.
- Block, Ned & Stalnaker, Robert, 1999. "Conceptual Analysis, dualism, and the explanatory gap", *Philosophical Review* 108:1-46.
- Block, N., 2007. Max Black's Objection to Mind-Body Identity, in T Alter and S Walter (eds) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and physicalism*, Oxford University Press.
- Bloom, P., 2004. *Descartes' Baby*, Basic Books.
- Brown, R., 2010. Deprioritizing the A Priori Arguments Against Physicalism, *Journal of Consciousness Studies*, 17:47-69.
- Carnap, R. (1950). 'Empiricism, semantics and ontology', *Revue Internationale de Philosophie* 4; reprinted in his *Meaning and necessity*.
- Carruthers, P., Veillet, B, 2007. The Phenomenal Concept Strategy, *Journal of Consciousness Studies* 14. Consciousness and the Laws of Physics.
- Carroll, Sean M. (2021). *Journal of Consciousness Studies* 28 (9-10):16-31.
- Chalmers, D, 1996. *The Conscious Mind*, New York: Oxford University Press.
- Chalmers, D., and Jackson, F., 2001. Conceptual Analysis and Reductive Explanation, *Phil Review* 110:315-61.
- Chalmers, D., 2002a. Does Conceivability Entail Possibility?, In (T. Gendler & J. Hawthorne, eds) *Conceivability and Possibility*, Oxford University Press, pp.145-200.
- Chalmers, D., 2002b. Consciousness and its place in nature. In (D. Chalmers, ed) *Philosophy of Mind: Classical and Contemporary Readings* (Oxford University Press, 2002). Also in (S. Stich & F., Warfield, eds) *Blackwell Guide to Philosophy of Mind*. Blackwell, 2003. Reprinted (abridged) in ( W. Lycan & J. Prinz, eds) *Mind and Cognition: A Reader* (Blackwell, 2007).
- Chalmers, D., 2003. The Content and Epistemology of Phenomenal Belief, in Q. Smith and A. Jokic (eds), *Consciousness: New Philosophical Perspectives*, Oxford University Press.

- Chalmers, D., 2004. The foundation of two-dimensional semantics, in *Two-Dimensional Semantics: Foundations and Applications*, Oxford University Press.
- Chalmers, D., 2006. Two-Dimensional Semantics, in *Oxford Handbook to the Philosophy of Language*. (Lepore and Smith, eds.), Oxford University Press.
- Chalmers, D., 2007. “Phenomenal Concepts and the Explanatory Gap.” in T Alter and S Walter (eds) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and physicalism*, Oxford University Press.
- Chalmers, D., 2009. “The Two-Dimensional Argument Against Materialism”, in B. McLaughlin (ed.) *The Oxford Handbook of the Philosophy of Mind*.
- Chalmers, D., 2017. Panpsychism and panprotopsychism, in *Panpsychism: Contemporary Perspectives*, Oxford University Press.
- Davidson, D., 1973. “Radical Interpretation”, in *Inquiries into Truth and Interpretation*, Oxford University Press, 1984.
- Dennett, D.C., 2016. Illusionism as the obvious default theory of consciousness, *Journal of Consciousness Studies*, 23 (11–12), pp. 65–72. Reprinted in Frankish, K. (ed.) (2017) *Illusionism as a Theory of Consciousness*, Exeter: Imprint Academic.
- Descartes, René (1641/1985). *Meditations*, in *The Philosophical Writings of Descartes*, vol. 2, trans. J. Cottingham et al., Cambridge: Cambridge University Press.
- Dretske, F., 1988. *Explaining Behavior: Reasons in a World of Causes*. Cambridge: MIT Press.
- Fodor, J., 1998. *Concepts: where cognitive science went wrong*, Oxford University Press.
- Fodor, J., 1990. *A Theory of Content and Other Essays*. Cambridge: MIT Press.
- Frankish, K., 2007. “The anti-zombie argument”. *Philosophical Quarterly*, 57 (229). pp. 650-666.
- Frankish, K., 2016. “Illusionism as a Theory of Consciousness, *Journal of Consciousness Studies* 23(11-12).
- Hill, C. and McLaughlin, B., 1999. There are Fewer Things Than Are Dreamt of in Chalmers's Philosophy, *Philosophy and Phenomenological Research* 59(2), 445-454.

- Glymour, C. (1970). Theoretical realism and theoretical equivalence. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*.
- Goff, Philip, 2017. *Consciousness and Fundamental Reality*, New York: Oxford University Press.
- Graziano, M., 2013. *Consciousness and the Social Brain*, New York: Oxford University Press.
- Humphrey, N., 2011. *Soul Dust: The Magic of Consciousness*, Princeton, NJ: Princeton University Press.
- Jackson, F., 1982. "Epiphenomenal Qualia", *Philosophical Quarterly* 32: 127-36.
- Jackson, F., 1993. Armchair Metaphysics. In *Philosophy in Mind*, ed. M. Michael and John O'Leary-Hawthorne. Dordrecht: Kluwer.
- Jackson, F., 2003. "Mind and illusion". In (A. O'Hear, ed) *Minds and Persons*. Cambridge University Press.
- Jackson, F., 2007. "A Priori Physicalism", in *Contemporary Debates in Philosophy of Mind*, Blackwell Publishing.
- Jackson, F., 2020. A Priori Physicalism, in U Kriegel (ed) *The Oxford Handbook of the Philosophy of Consciousness*, Oxford University Press.
- Kaplan, D, 1968-1969. "Quantifying In," *Synthese*, 19.1-2 (Dec.): 178-214.
- Kirk, R. 2005. *Zombies and Consciousness*, Oxford University Press.
- Kriegel, U., 2004. "Consciousness and Self-Consciousness". *Monist*: 182-205.
- Kripke, S., 1972. *Naming and Necessity*. Cambridge: Harvard University Press.
- Levin, J., 2007. "What is a phenomenal concept?", in T Alter and S Walter (eds) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and physicalism*, Oxford University Press.
- Levine, J, 2001. *Purple Haze*, Oxford University Press.
- Levine, J. 2006. "Conscious awareness and (self)representation", in *Self-Representational Approaches to Consciousness*, Uriah Kriegel and Kenneth Williford (eds.), MIT/Bradford.

- Levine, J., 2007. "Phenomenal concepts and the materialist constraint", in T Alter and S Walter (eds) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and physicalism*, Oxford University Press, 145-165..
- Lewis, D., 1966. "An argument for the Identity Theory", *Journal of Philosophy*, LXIII. 1 (January 6, 1966): 17-25.
- Lewis, D., 1983a. New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61:343-77.
- Lewis, D., 1983b. "Radical Interpretation", in *Philosophical Papers*, vol 1. Oxford University Press.
- Loar, B. (1990/1997). "Phenomenal states", *Philosophical Perspectives* 4, Action Theory and Philosophy of Mind, 81-108, published again with revisions in *The Nature of Consciousness*, (Block, Flanagan, Güzeldere, eds.), MIT Press, 597-616.
- Loewer, B, 1978. "Leibniz and the Ontological Argument", *Philosophical Studies* 34 (1978) 105-109.
- Loewer, B., 1995. "An Argument for Strong Supervenience" In (E. Savellos & U. Yalcin, eds.) *Supervenience: New Essays*. Cambridge University Press.
- Loewer, B, 2020. The Package Deal Account of Laws and Properties. *Synthese* 199 (1-2):1065-1089.
- Marton, Peter, 1998. "Zombies versus Materialists: The Battle for Conceivability", *Southwest Philosophy Review*, **14**: 131-138.
- McLaughlin, B, 2007. "On the limits of A Priori Physicalism", *Contemporary Debates in Philosophy of Mind*, Blackwell Publishing.
- Melnyk, A., 2003. *A Physicalist Manifesto: Thoroughly Modern Materialism*, Cambridge University Press, Cambridge.
- Miller, Kristie, (2005). The metaphysical equivalence of three and four Dimensionalism, *Erkenntnis* 62: 91–117.
- Miller, Kristie, (2005). What is metaphysical equivalence?, *Philosophical papers*, Vol. 34, No 1. 44-74.
- Miller, Kristie (2017) A hyperintensional account of metaphysical equivalence, *The Philosophical Quarterly* Vol. 67, No. 269.

- Millikan, R., 1989. "Biosemantics." *Journal of Philosophy* 86: 281-97.
- Mørch, Hedda Hassel, 2019. "The Argument for Panpsychism from Experience of Causation", in Seager, W.E. *Routledge Handbook of Panpsychism*, London, New York: Routledge.
- Nagel, T., 1974, "What Is It Like to Be a Bat?", *Philosophical Review* 83:435-50.
- Nida-Rümelin, M., 2007. "Grasping Phenomenal Properties", in T Alter and S Walter (eds) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and physicalism*, Oxford University Press.
- (2017). "Realism about Identity and Individuality of Conscious Beings", in Christian Kanzian, Sebastian Kletzl, Josef Mitterer, Katharina Neges (eds.), *Realism - Relativism - Constructivism: Proceedings of the 38th International Wittgenstein Symposium in Kirchberg*, Berlin, Boston: De Gruyter, 279–292.
- Papineau, David. 1993. *Philosophical Naturalism*. Oxford: Blackwell.
- Papineau, D., 1995. "Arguments for Supervenience and Physical Realization." In (E. Savellos & U. Yalcin, eds.) *Supervenience: New Essays*. Cambridge University Press.
- Papineau, D., 2002. *Thinking about consciousness*, Oxford UP.
- Papineau, D., 2007a. "Phenomenal and Perceptual Concepts", in T Alter and S Walter (eds) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and physicalism*, Oxford University Press.
- Papineau, D., 2007b. "Kripke's Proof Is Ad Hominem Not Two-Dimensional", *Philosophical Perspectives*, Vol. 21, Philosophy of Mind, pp. 475-494.
- Papineau, D., 2020. "The problem of consciousness", in U Kriegel (ed) *The Oxford Handbook of the Philosophy of Consciousness*, Oxford University Press, 17-36.
- Pautz, Adam, 2020. "How to Achieve the Physicalist Dream Theory of Consciousness: Identity or Grounding?", In G. Rabin (ed.), *Grounding and Consciousness* (forthcoming).
- Perebrom, D. (2011) *Consciousness and the Prospects of Physicalism*, Oxford: Oxford University



Press.

Perry, J., 2001. *Knowledge, Possibility and Consciousness*. MIT Press.

Piccinini, G. 2015. Access Denied to Zombies, *Topoi*, 10.1007/s11245-015-9323-6.

Putnam, H. 1963. An Examination of Gröbaum's Philosophy of Geometry, in *Philosophy of Science. The Delaware Seminar*, Vol. 2 (ed. by B. Baumrin).

Putnam, H., 1982, Why There Isn't a Read-Made World? *Synthese*, Vol. 51, No. 2, Realism, Part I, 141-167

Putnam, H., 1989. *Representation and Reality*, Cambridge, Massachusetts: The MIT Press.

Quine, W.V., 1975. On Empirically Equivalent Systems of the World, *Erkenntnis* 9(3), pp. 313-328.

Reichenbach, H. 1961. *Experience and Prediction*, University of Chicago Press.

Rey, G., 2007. "Phenomenal Content and the Richness and Determinacy of Color Experience", *Journal of Consciousness Studies*, 14.

Robinson, H., 1993. "The Anti-materialist Strategy and the Knowledge Argument", in *Objections to physicalism*, ed. Howard Robinson. Oxford: Oxford University Press.

Salmon, W. 1966. Verifiability and Logic, in *Mind, Matter and Method* (ed. by P. Feyerabend and G. Maxwell).

Schaffer, Jonathan, 2017. The Ground Between the Gaps, *Philosophers' Imprint* 17.

Searle, John. 1992. *The Rediscovery of the Mind*. Cambridge: MIT Press.

Sider, Theodore, 2011. *Writing the Book of the World*. Oxford:Oxford University Press.

Sider, Theodore, 2016. "TheoreticalEquivalence." Presented as Lecture 6 of the 2016 Locke Lectures, Oxford University.

Soames, S, 2004. Reference and Description: The Case against Two-Dimensionalism, Princeton University Press.

- Stoljar, D., 2001. Two conceptions of the physical, *Philosophy and Phenomenological Research* 62 (2):253-81.
- Stoljar, D., 2005. "Physicalism and phenomenal concepts", *Mind and Language*.
- Sturgeon, S., 2000. *Matters of Mind: Consciousness, Reason and Nature*. Routledge.
- Strawson, G., 2006. Realistic monism: why physicalism entails panpsychism, in *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?*, A. Freeman (ed.), Exeter: Imprint Academic.
- Tye, M., 2003. "A theory of phenomenal concepts", *Philosophy*.
- Van Fraassen, B., 1980. *The Scientific Image*, Oxford: Oxford University Press.
- White, S., 2007. "The Argument for the Semantic Premise", in T Alter and S Walter (eds) *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and physicalism*, Oxford University Press.
- Williams J. R. G., 2008. Ontic Vagueness and Metaphysical Indeterminacy, in *Philosophy Compass* 3/4: 763–788, 10.1111/j.1747-9991.2008.00151.x
- Yablo, S., 1993. "Is Conceivability a Guide to Possibility?" *Philosophy & Phenomenological Research* 53 pp. 1-42.
- Yablo, S. (1999). Concepts and Consciousness. *Philosophy and Phenomenological Research*, 59(2), 455–463.
- Yablo, S., 2002. "Textbook Kripkeanism and the Open Texture of Concepts", *Pacific Philosophical Quarterly*, Vol. 81: 1, pp. 98-122.