

## **Powers and the mind-body problem\***

(forthcoming in *International Journal of Philosophical Studies*)

István Aranyosi

*The paper proposes a new line of attack on the conceivability argument for mind-body property dualism, based on the causal account of properties, according to which properties have their conditional powers essentially. It is argued that the epistemic possibility of physical but not phenomenal duplicates of actuality is identical to a metaphysical (understood as broadly logical) possibility, but irrelevant for establishing the falsity of physicalism. The proposed attack is in many ways inspired by a standard, broadly Kripkean approach to epistemic and metaphysical modality.*

The most discussed arguments for mental-physical property dualism in recent years have been the so-called conceivability arguments. Recently, the most sophisticated champion of such arguments has been David Chalmers. His recipe for such an argument goes like this<sup>1</sup>. Take two actual truths, P and Q, such that most of us have the intuition that ‘P → Q’ is not *a priori*, and so ‘P & ¬Q’ is conceivable. Then argue that the right kind of

---

\* Acknowledgments: thanks to David Chalmers, Frank Jackson, Daniel Stoljar, Stephan Leuenberger, Jessica Wilson, Laurie Paul, Jonathan Schaffer, Robert Adams, Alex Byrne, Christopher Hill and two anonymous referees for valuable feedback.

<sup>1</sup> ‘The two-dimensional argument against materialism’. In B. McLaughlin & A. Beckermann, eds) *Oxford Handbook of the Philosophy of Mind*. Oxford University Press (forthcoming). ‘Consciousness and its place in nature’. In S. Stich & T. Warfield (eds.) *Blackwell Guide to Philosophy of Mind*. Blackwell, 2003.

conceivability entails a certain kind of possibility. Then argue that the conceivability of 'P &  $\neg$ Q' is the right kind to entail that kind of possibility of 'P &  $\neg$ Q'. Then argue that this latter possibility is the right kind of possibility to entail that the entities figuring in P and Q, respectively, are actually distinct.

This recipe yields property dualism for: 'the totality of physical truths' as a replacement for P, 'a phenomenal truth' as a replacement for Q, 'primary ideal positive conceivability' as replacement for 'the right kind of conceivability', and, finally, 'primary possibility' as a replacement for 'the right kind of possibility'. The meaning of these quoted phrases will be explained in due course.

A more popular name for the argument is 'the argument from the conceivability of zombies', zombies being physical duplicates of us, conscious beings, but lacking conscious experience whatsoever. So it is physicalists who are mainly interested in showing why the argument fails since according to physicalism a physical duplicate of the actual world is a duplicate of it in any other respect. I will make the assumption that "physical" means described/explained by physics.

I don't think the argument works. My motivation for this claim comes from the fact that I happen to subscribe to a certain version of the view that properties are to be individuated via the powers they bestow upon the objects that possess them, which view yields –as I will try to show here– a good reply to the conceivability argument.

In what follows, I will (1) expound the zombie argument, the replies that have been considered so far in the literature, and why those replies are considered unsatisfactory by the fans of the zombie argument, (2) expound the causal account of properties that I favour, (3) explain the way in which this account constitutes a basically

Kripkean reply to the argument and why it is better than the others, (4) discuss some worries that arise in connection with the reply, and, finally, (5) draw, in light of the previous points, some consequences regarding the way power essentialism could cope with some objections. In other words, I will offer both an application of the powers based conception of properties to the mind-body problem and, at the same time, a defence of this conception, based precisely on some ideas that will emerge in the context of applying it.

### **1. The argument**

Let me first introduce some notions that I will use when formulating the zombie argument and the reply.

*Metaphysical and ideal conceptual possibility.* Metaphysical possibility will be used in the sense of broadly logical possibility, which is the standard and traditional sense, which presupposes a single space of worlds, interpreted predicates, and the same strength for logical and metaphysical necessity<sup>2</sup>. So, for example, not only propositions like *that P and ¬P* will come out as impossible, but also propositions like *that there is a round square*, and under the assumption that “water” is a rigid term, *that water is not a chemical compound*. By ideal conceptual possibility, on the other hand, I will mean broadly logical possibility consistent with the set of all *a priori* knowable propositions. So, for instance, given the set K of all *a priori* knowable propositions, P will be ideally conceptually impossible if and only if K contains a set of propositions P\* such that P &

---

<sup>2</sup> See Saul A. Kripke, *Naming and Necessity*, Harvard University Press, 1980 [1972], pp. 35-6, and Sydney Shoemaker, “Causal and metaphysical necessity”, *Pacific Philosophical Quarterly*, Volume 79, Number 1, March 1998, p. 60.

$P^*$  is metaphysically (i.e. broadly logically) impossible. So ideal conceptual possibility is the same as the set of propositions not ruled out *a priori*, i.e. whose negations are not entailed by what is knowable *a priori*.

*Broad Kripkeanism.* What I will call 'a broadly Kripkean account' of cases when there is a gap between ideal conceptual possibility and metaphysical possibility, i.e. a modal illusion, is the following. When a proposition  $S$  is *a posteriori* necessary, its negation is both metaphysically impossible and ideally conceptually possible; furthermore, the ideal conceptual possibility of  $S$  is the same as the misdescribed metaphysical possibility of some other proposition  $S^*$ , distinct from  $S$ . Take, for instance, the textbook example *Water is  $H_2O$* . *Water is not  $H_2O$*  is metaphysically impossible and ideally conceptually possible, given the *a priori* truths about water and  $H_2O$ ; furthermore, the ideal conceptual possibility of *Water is not  $H_2O$*  is the same as the misdescribed metaphysical possibility of something that superficially resembles water and is not composed of  $H_2O$  molecules.

*Primary and secondary intensions.* The above ideas related to ideal conceptual possibility, metaphysical possibility, and broad Kripkeanism, are nicely accounted for by Chalmers' two-dimensional (2D) semantics, by way of a distinction between the *a priori* aspect and the *a posteriori* aspect of the meaning of concepts. The primary intension (1-intension) of a concept is the information speakers *a priori* associate with that concept. For instance, the 1-intension of the term "water" is thought to be represented by way of a non-rigid descriptive phrase, more or less like "the transparent, odourless, colourless, drinkable liquid, that falls as rain, flows in rivers, and covers most of our planet." The secondary intension (2-intension) is the *a posteriori* aspect of the meaning of a term - the

same as the actual reference of that term. For instance, in the case of the term "water", the 2-intension is "H<sub>2</sub>O", supposing that indeed water is actually composed of H<sub>2</sub>O molecules. Ideal conceptual possibility, is then understood in this paper as truth of the 1-intension at some possible world. Metaphysical possibility will be understood as truth of the 2-intension at some possible world. Possible worlds will be understood as broadly logically possible worlds, and, furthermore, I will assume that there is only one modal space, the space of broadly logically possible worlds<sup>3</sup>. So ideal conceptual possibility and metaphysical possibility will be referred to by '1-possibility' and '2-possibility', respectively.

What I have called 'broad Kripkeanism' about modal illusions is then the following, in Chalmers' terminology: when P is *a posteriori* necessary,  $\neg P$  is 1-possible and 2-impossible; further, the 1-possibility of  $\neg P$  implies that there is a P\* ( $P \neq P^*$ ) such that  $\neg P^*$  is 2-possible.

*Conceivability.* Conceivability will be regarded as a way to reliably access the modal space. Corresponding to 1- and 2-possibility we will have 1- and 2-conceivability, as ways of reliably accessing 1- and 2-possibility, respectively. Following Chalmers<sup>4</sup>, let us mention two more dimensions of analysis for conceivability: the ideal/non-ideal dimension and the positive/negative dimension.

A proposition is ideally conceivable iff no amount of *a priori* reasoning can rule out that proposition.

---

<sup>3</sup> As in Frank Jackson, *From Metaphysics to Ethics. A defence of conceptual analysis*. Oxford University Press, 1998, p. 70.

<sup>4</sup> "Does conceivability entail possibility", sections 1 and 2.

A proposition is positively conceivable iff it is imaginable. And a proposition is negatively conceivable iff it is consistently supposable. It is thought in the current literature on modal epistemology that supposition is less restrictive than imagination when it comes to accessing the realm of modal truths. For instance, even overt contradictions and negations of trivial truths can be supposed for the purpose of a *reductio*, while even things like a round square are unimaginable. Negative conceivability, however, is supposition without the derivability of a contradiction.

This being said, let us formulate the conceivability argument. For P understood as the totality of physical truths and Q as a phenomenal truth, I will assume that P & ¬Q is ideally and positively conceivable; I make this assumption because the proposed reply to the argument will be based exclusively on issues related to 1- and 2-conceivability and possibility – the traditionally Kripkean niche of attack on the argument.

Given that to show that some scenario is metaphysically possible is to show that it is 2-possible, the purported anti-physicalist conclusion of the zombie argument will have to follow from the proposition that zombies are 2-possible, i.e. that there can be a world that duplicates the 2-intension of all the physical truths about actual objects and properties but fails to duplicate the 2-intension of actual phenomenal truths.

The crucial point in Chalmers' argument is, in fact, that once we accept that zombies are 1-possible, we are driven to a disjunction of two propositions, both of which are equivalent to the falsity of physicalism. First, we should accept the 1-possibility of zombies, if we accept their 1-conceivability. Second, if phenomenal concepts do have distinct 1- and 2-intension, each concepts 1-intension can be construed as the 2-intension of some other phenomenal concept. Then everything turns out to depend on whether the

physical truths have distinct 1- and 2-intension or not, from which we get the following disjunction of propositions:

- (a) On the assumption of coinciding 1- and 2-intensions for P:  $P \ \& \ \neg Q$  is both 1- and 2-possible, which is equivalent to physicalism being false.
- (b) On the assumption of distinct 1- and 2-intensions for P:  $P \ \& \ \neg Q$  is 1-possible and 2-impossible, which is equivalent to “panprotopsychism” being true, where panprotopsychism is defined as the view according to which there are truths not accounted for by physical theory which collectively entail both the physical truths and the phenomenal ones.

We can now express the argument more formally:

#### *Conceivability of Zombies Argument*

1.  $P \ \& \ \neg Q$  is 1-conceivable. (Premise: conceivability intuition)
2. 1- conceivability of R entails 1- possibility of R. (Premise: Conceivability-Possibility Principle)
3.  $P \ \& \ \neg Q$  is 1-possible. (From 1, 2)
4.  $P \ \& \ \neg Q$  is either 2-possible or not. (Tautology)
5. If  $P \ \& \ \neg Q$  is 2-possible, then physicalism is false. (By definition of ‘physicalism’)
6. If  $P \ \& \ \neg Q$  is 1-possible but not 2-possible, then either (a) P has distinct 1- and 2-intension but Q does not, or (b) Q has distinct such intensions but P does not, or (c) both have distinct such intensions. (From the assumption of compositionality of 2D semantics)

7. If  $P \ \& \ \neg Q$  is 1-possible but not 2-possible and (a), then physicalism is false and panprotopsychism is true. (By definition of 'physicalism' and of 'panprotopsychism')
8. If  $P \ \& \ \neg Q$  is 1-possible but not 2-possible and (b), then physicalism is false. (By definition of 'physicalism' and the assumption that the 1-intension of a phenomenal concept is identical to the 2-intension of some other phenomenal concept)
9. If  $P \ \& \ \neg Q$  is 1-possible but not 2-possible and (c), then physicalism is false and panprotopsychism is true. (By definition of 'physicalism' and 'panprotopsychism', and the assumption that the 1-intension of a phenomenal concept is identical to the 2-intension of some other phenomenal concept)
10. Physicalism is false. (From 5, 6, 7, 8, 9)

There have been two main replies to the argument: Type A (i.e. *a priori*) physicalism, which denies premise 1, and Type B (i.e. *a posteriori*) physicalism, which denies (or has to deny according to Chalmers) premise 2.

Type A physicalism, as it is usually understood<sup>5</sup>, is in fact an overreaction to premise 1 since it denies even the *prima facie* conceivability of zombies, not only their ideal conceivability. Supposing we do have a conceivability intuition regarding zombies, this kind of *a priori* physicalism is ruled out.

Type B physicalism is unconvincing because there is no standard Kripkean account of an *a posteriori* necessary statement S such that its negation is primarily

---

<sup>5</sup> Or at least as it is defined by Chalmers in his 'Consciousness and its place in nature'.



conceivable, but not primarily possible – what Chalmers has called a ‘strong necessity’. There are no other examples of strong necessities, so appeal to them in the context of the mind-body problem has no motivation other than saving physicalism.

Chalmers is happy, as reflected by propositions 7 and 9, with the idea that P has distinct 1- and 2-intensions, and so  $P \ \& \ \neg Q$  is not 2-possible after all. This is what I have earlier mentioned as panprotopsychism - the ontological view according to which both physical and phenomenal properties are entailed by some more fundamental ones. An alternative name for this doctrine has been proposed by Daniel Stoljar<sup>6</sup> – *object physicalism* – understood and defended as a kind of non-standard physicalism, based on the existence of intrinsic, categorical properties that form the supervenience base of properties that physical theory describes, a base whose nature we are currently ignorant about. As opposed to this, standard physicalism –which Stoljar calls *theory physicalism* - is based on a supervenience base consisting of precisely the properties that physics describes. Obviously, what I mean by ‘physicalism’ in the above argument is not the same as object physicalism, and is roughly what Stoljar means by theory physicalism.

## **2. The causal account of properties**

The reply to the zombie argument I will shortly expound is based on the causal account of properties I favour, which is inspired by Sydney Shoemaker’s more recent such theory,<sup>7</sup> and consists of three theses.

---

<sup>6</sup> ‘Two conceptions of the physical’, *Philosophy and Phenomenological Research*, Vol. LXII, No.2, March 2001, 253-281.

<sup>7</sup> ‘Causal and metaphysical necessity’, *Pacific Philosophical Quarterly*, Volume 79, Number 1, March 1998, 59-77.

I first define ‘conditional powers’ of properties, where causal relata are assumed to be events, understood as ordered triples of objects, properties and times (times are omitted in the definitions):

**Forward-looking powers:** A property  $F$  has/bestows a conditional power  $P(\mathbf{A})$  to cause  $e$  iff there is a set of properties  $\mathbf{A}$ , such that for any individual  $x$ , if the elements of  $\mathbf{A}$  are instantiated,  $x$ -having- $F$  causes  $e$ .

**Backward-looking powers:** A property  $F$  has/bestows a conditional power  $P(\mathbf{A})$  to be caused by  $c$  iff there is a set of properties  $\mathbf{A}$ , such that for any individual  $x$ , if the elements of  $\mathbf{A}$  are instantiated,  $x$ -having- $F$  is caused by  $c$ .

For some property  $F$  and conditional power  $P(\mathbf{A})$ , I will call set  $\mathbf{A}$ , the *condition set*, and  $e$  and  $c$  the *forward-looking manifestation* and the *backward-looking manifestation* of  $F$ , respectively. I then state the three theses.

***Individuation Thesis (Trans-world):*** Both forward-looking and backward-looking conditional powers of a property  $F$  are essential to its being instantiated. More formally:

---

For any property  $F$ , for the set  $\mathbf{P}$  of all its conditional powers, and for all manifestations  $M$  in a world  $W$  of any  $P(\mathbf{A}) \in \mathbf{P}$ , given some condition set  $\mathbf{A}$ , if the members of  $\mathbf{A}$  are instantiated at some world  $W^* \neq W$ , and some conditional power  $P(\mathbf{A}) \in \mathbf{P}$  is manifested as  $M^* \neq M$  in  $W^*$ , then  $M^*$  is the manifestation of a property  $F^* \neq F$ .

**Identity Thesis (Intra-world):** For any world  $W$ , and any two properties  $F$  and  $G$  instantiated in  $W$ ,  $F = G$  if and only if the totality of manifestations of  $F$  in  $W$  is identical with the totality of manifestations of  $G$  in  $W$ .

**Existence Thesis ([my version of] The Eleatic Principle or Alexander's Dictum):** (a) If a property exists, then it has some conditional powers, and (b) if a property is instantiated, then it is causally manifested, either forward or backward, in all worlds in which *some*<sup>8</sup> members of its condition set are instantiated.

---

<sup>8</sup> As pointed out by a referee, point (b) of the Existence Thesis appears to be, at first sight, slightly at odds with the Individuation Thesis. But the reference to *some* of the members of the condition set in the former versus the reference to *all* of them in the latter makes the difference. What the former says is that by varying the number of properties that are instantiated, which are members of the condition set, one does not fail to have the relevant property instantiated. For example, suppose that there are two actual conditions for a billiards ball with certain properties to have caused another one to move with a certain velocity: that the billiards table was straight and that the ball hit the other one in the middle. There is no reason to think that in a world where the table is not straight, or in one in which the ball does not hit the other one exactly in the middle the property of the second ball moving with the same velocity is not instantiated. By contrast, what the Individuation Thesis asserts is that in worlds in which *all* the members of the condition set are instantiated, i.e. when we don't vary the number of the instantiated

According to this view, conditional powers are essential to, but not exhaustive of what a property is. It is also worth emphasising that this essentiality rather than exhaustiveness is to be understood to entail that the Individuation Thesis will sanction true statements relating a property and any subset of its conditional powers as *a posteriori* and necessary, rather than *a priori*.

### 3. The reply

First, if the Individuation Thesis is true, then zombies are impossible. In the alleged zombie world, on the supposition (a) that phenomenal properties are actually causally responsible for bringing about some physical effects, namely, behaviour, the backward-looking manifestations of the latter (i.e. its causes) are different from its actual backward-looking manifestations. This is so because there are no phenomenal properties instantiated, and so, contrary to the supposition, the alleged zombie world is not a behavioural duplicate of actuality. On the supposition (b) that phenomenal properties are at least caused in the actual world by the instantiation of some physical properties, namely, the stimuli, the forward-looking manifestations of the latter (i.e. its effects) are different from their actual forward-looking manifestations. This is so because there are no phenomenal properties instantiated, and so, contrary to the supposition, the alleged zombie world is not a stimulus duplicate of actuality. So  $P \ \& \ \neg Q$  is 2-impossible.

---

members of the condition set, the property we focus on must have the same causes and effects from one world to another. Using our billiards example: in all worlds that share with actuality the two conditions –straight table and the angle of impact- the effects and the causes of the billiards ball with its actual features must be the same from one world to another.

The question is then this: is  $P \ \& \ \neg Q$  both 1- and 2- impossible, or only 2- impossible (and 1- possible)? In other words, are zombies both *a priori* and *a posteriori* impossible or only *a posteriori* impossible (and *a priori* possible)? Before approaching this question let me put forward two remarks.

*Remark 1:* the supposition in point (b), together with the Existence Thesis, is compatible with epiphenomenalism about phenomenal properties, since it requires only that they be part of the causal structure of the world, not necessarily as causes, but as effects. So what we have just established does not beg any question against any view about the place and role of phenomenal properties in the causal web of the world. (Contrast with Jaegwon Kim's version of Alexander's Dictum, where a property does need forward-looking manifestations in order to be "real"<sup>9</sup>).

*Remark 2:* my version of the Existence Thesis is or should be more or less uncontroversial, unlike many other formulations. In particular, it is far from what might be called *metaphysical naturalism*, the view that a property exists insofar as it is actually causally manifested. This view is rightly regarded as based on prejudice<sup>10</sup>. My version of the thesis only requires that there be *possible* condition sets such that a particular instantiating the members of that set will have a causal manifestation if it also instantiates the property.

---

<sup>9</sup> *Supervenience and Mind*. Cambridge University Press, Cambridge, 1993, p. 348.

<sup>10</sup> Oliver, Alex, The metaphysics of properties, *Mind*, Vol. 105, 417, January 1996, 1-80, p. 8.

Going back to our question, note first that the Individuation Thesis is compatible with  $P \ \& \ \neg Q$  being 1- conceivable, since, as I have mentioned, the thesis yields statements relating properties and their causal powers as necessary, but *a posteriori*, and *a posteriori* truths, even when necessary are 1-conceivably false. Further, according to the approach of broad Kripkeanism that I put forward such statements have ideally epistemically possible negations, i.e. their negations are 1-possible. So the answer to our question is that  $P \ \& \ \neg Q$  is 2-impossible, but 1-possible.

The 1-possibility of zombies is just what the fans of the zombie argument are after, but note that the only way  $P \ \& \ \neg Q$  could be 2-impossible but 1-possible, if our account of properties is right, is for the 1-intension of predicate terms in P and Q to involve no reference to conditional powers or involve reference to some but not all the existing conditional powers<sup>11</sup>. Note further that in the zombie scenario all the members of the *physical condition set* are instantiated, therefore, all the actual physical manifestations are supposed to be present there. In other words, all the properties accounted for by physics are supposed to be instantiated, according to the conceived scenario. But if the 1-intension of predicate terms in P and Q involves no reference to conditional powers or involves reference to some but not all the conditional powers, and, at the same time, it does involve reference to manifestations, it has to be the case that the 1-intension fails to involve reference to at least some of those manifestations *qua causal manifestations of powers*. In other words, there is a coherent interpretation according to which the zombie world involves complete duplication of actual world physical regularities, but either

---

<sup>11</sup> The latter - when the 1-intensions of P and Q involves reference to some but not all the conditional powers - would be the case if one wanted, in the imagined world, to keep everything free of causal gaps *except* regions where phenomenal properties are supposed to be instantiated.

contains no causation whatsoever or contains some but not all causation that goes on in actuality, where causation is understood as non-Humean, i.e. as a necessary connection. As I like to put it, according to this understanding, the zombie world is a physical mirror image of our world, but the mirror reflecting it is a broken one.

But how could this be the case, i.e. conceivable, since it contradicts point (b) of the Existence Thesis? Indeed, if this is what the 1-possibility of  $P \ \& \ \neg Q$  amounts to then it seems that it is a world in which, contrary to the causal account, there are instantiated alien properties with no forward or backward manifestations. However, this is too quick; there is one more option in logical space that could account for the situation. Hume-worlds – worlds with no causation but only mere constant conjunction of events - are 1-conceivable and 1-possible under our causal account of properties, if the notion of a property is itself two-dimensional, with different 1- and 2-intensions. Since according to the theory such Hume-worlds are 2-impossible, this means that the 1-possibility of the zombie world is the 2-possibility of a world either with no properties whatsoever, but only items that *seem to be* properties, call them ‘schmoperties’, or a world containing such schmoperty regions. In any of these cases not all physical properties of the actual world are instantiated.

Where does this leave us? The metaphysical possibility of a world that does not duplicate all the physical properties and in which there is no phenomenal consciousness is irrelevant to the truth of physicalism. What is relevant to physicalism is a world with the same physical properties but no phenomenal consciousness. Of course, if the only alternative for the 1-possibility of the zombie world were the 2-possibility of a world with all the actually instantiated properties being instantiated *except* the phenomenal ones,

such that those instantiated properties only seem physical, then that world would be a world that lacks something more fundamental than what we take to be physical in our world, therefore, panprotopsychism would be true. However, given our account of properties and what has just been derived above, this is not [and cannot be] the case.

In sum, I deny, based on the epistemic possibility of my account of properties being true, propositions 7, 8, and 9, by putting forward an alternative for P's 1- and 2-intensions to diverge. The conclusion then is: the zombie argument fails because the 1-possibility of  $P \ \& \ \neg Q$  does not entail the falsity of physicalism. Importantly, the motivation for this result is broadly Kripkean, which means that, contrary to what Chalmers has argued, *there is* a standard broadly Kripkean way to block the argument. To my knowledge, this niche – the '*schmoperty move*' - has not been identified and therefore exploited so far. Indeed, in a recent article, Chalmers thinks that the appeal to a Kripke-inspired metaphysical necessity of natural laws does not have any impact on the zombie argument, *given his assumption that 1-possible zombies are enough for the argument to work (and strong necessities are therefore needed to block it)*, except on a strong reading of law necessitarianism, which itself presupposes 'un-Kripkean' (read: 'unjustifiable by Kripkean considerations') strong necessities:

Some philosophers hold that the laws of nature are metaphysically necessary. On some views of this sort [...], this necessity arises for broadly Kripkean reasons: the reference of terms such as "mass" is fixed *a posteriori* to a certain very specific property, so that worlds with different laws do not contain mass. I think this view is



implausible, but in any case it is compatible with an entailment from primary conceivability to primary possibility. If G' is a counterfactual statement [...], then G' is both primarily conceivable and primarily possible. G' is verified by a metaphysically possible world W considered as actual, although not by W considered as counterfactual. (Considered as counterfactual, W contains "schmass", not mass.) So there are no strong necessities here.

There is a stronger view on which the laws of every world are exhausted by actual-world laws, applying to actual-world properties. On this sort of view, even "schmass" worlds are metaphysically impossible: G' will be primarily conceivable but not even primarily possible. On this view, laws of nature are strong necessities. There is no reason to accept this view, however. [...] Proponents of necessary laws usually appeal to Kripke's necessary *a posteriori* for support, but the Kripkean cases support at best the weak view in the previous paragraph.<sup>12</sup>

I want to emphasise at this point that the move is Kripkean in spirit and it shows precisely that Chalmers' assumption that the 1-possibility of zombies is sufficient for the argument to prove the falsity of physicalism is not warranted; and I don't appeal to anything like strong necessities. Importantly, to the extent that my reply can be regarded as compatible with type B physicalism (see the first objection in the next section), it shows not what

---

<sup>12</sup> "Does conceivability entail possibility?", p. 190.

type B physicalism is currently (and wrongly!) supposed to be forced to be committed to, but rather what type B physicalism *should be committed to*, namely, the denial of the passage from the 1-possibility of zombies to the falsity of physicalism.<sup>13</sup>

Finally, phenomenal inversion, partial zombies (e.g. visual, but not, say, auditory), or any abnormal (i.e. not as in the actual world) phenomenal distribution are all 2-impossible. Then we can run the same argument as above against all these cases: all will turn out to be 1-conceivable and 1-possible, but their 1-possibility is irrelevant to the truth of physicalism.

#### 4. Discussion

Let me discuss a few worries that arise in connection with the schmoperty move against the zombie argument.

The first concerns the reliance on the Individuation Thesis, and contends that my reply has to presuppose property dualism, so even if it is successful against the argument, it is so in virtue of a dualistic picture that is assumed at the outset, and so, at least, it is not

---

<sup>13</sup> A corollary is that according to what type B physicalism should be, zombies are *not 2-conceivable*. This means that there really is an *a priori* element even to a *posteriori* physicalism. This shouldn't be surprising at all in light of what I have been calling broad Kripkeanism, because according to it the necessity bit of all standard Kripkean examples of a *posteriori* necessity stems from a *a priori* necessity at the level of propositions, the aposteriority bit stemming from the linguistic representation of the truths embodied by those propositions. For example, the necessity of 'water is H<sub>2</sub>O' on the assumption of rigid designation stems from the necessity of the proposition *H<sub>2</sub>O is H<sub>2</sub>O*, while its aposteriority stems from the descriptions associated with the terms 'water' and 'H<sub>2</sub>O'. Frank Jackson has pointed out to me that he has recently independently arrived at basically the same idea about what a *posteriori* physicalism should be committed to. See his 'On ensuring that physicalism is not a dual attribute theory in sheep's clothing', *Philosophical Studies* Vol. 131, No. 1, 227-249, October 2006, where he argues that the aposteriority bit in a *posteriori* physicalism cannot be *de re* (about properties), but only *de dicto* (about predicates), on pain of an unwanted commitment to necessitarian attribute dualism.

usable by the physicalist. In particular, the reason why the physical properties that are supposed to be instantiated in the zombie world are not and cannot be the same as those instantiated in actuality is that the putative instantiation of some of the former is not caused by and/or not causing the instantiation of *phenomenal properties*. So the correct qualitative characterisation/description of the actual world, even that of its physical regions requires, according to the reply, qualitative characterisation of its phenomenal regions, and this is far from how we understand physical characterisation, namely, as not having to involve anything mental.

In reply suppose first that I *do* have to assume a dualist perspective from which to deploy the analytical munition of the powers-based account against the zombie argument. Does this change the fact that if the powers-based account of properties is true, the 1-possibility of zombies and, more generally, any 1-possibility involved by property concepts according to which properties are duplicated without their causal essence being duplicated, comes out as irrelevant for the intended conclusion about actual property distinctness? I don't think so. Supposing I have assumed property dualism, what is the alternative starting point to this alleged dualist presupposition if the above conclusion above regarding the 1-possibility of zombies in particular is correct and if one wants to avoid it? The alternative seems to be that phenomenal properties are actually completely causally disconnected from the physical world. The alternative is disturbing twice over: first, it is itself a dualist starting point, so it is no better than what my starting point allegedly is; second, it is not even epistemically consistent with either epiphenomenalism about the phenomenal - understood as the conjunction of physical causal influence upon the phenomenal and phenomenal causal inefficacy – or with interactionism. It is only

consistent with parallelism, that is, the view according to which in the actual world the mental and the physical realms are independent and causally disconnected from each other. Now it is right that the *actual* truth of my account of properties does exclude parallelism, on the condition that both physical and phenomenal properties are actually instantiated and the reply is to work. But all we need for the reply to the zombie argument is the *epistemic possibility* of the account. The reply amounts to denying that premises 7, 8, and 9 exhaust the space of alternatives for the 1-possibility of zombies to be distinct from their 2-possibility; in effect it points to an alternative such that it makes that 1-possibility not only distinct but also irrelevant to the 2-possibility of zombies. Consequently, in order for the account not to have effect on the zombie argument it has to be *a priori* excluded, and that is what the alternative, parallelist picture has to do. But that is I think a question-begging way to save the zombie argument.

Second, and more importantly, I don't think I do have to assume a dualist picture in order for the reply to work. All that we start from is different predicates for actual phenomenal versus physical properties, without committing ourselves to either phenomenal/physical coreferentiality or non-coreferentiality. Then we entertain the epistemic possibility that these properties be part of the causal web of the world. This entertainability requirement, again, does not presuppose anything about the coreferentiality or non-coreferentiality of the predicates. Then we observe what follows from this possibility and the powers-based conception: the irrelevance of 1-possibility to the truth or falsity of physicalism. At no point is it forced upon us that the requirement that properties have forward- and/or backward-looking conditional powers essentially entail anything about their identity or distinctness; indeed, all we know is that causal

features have to be differently manifested in the conceived zombie world than they actually are, and that can be either because some interactionist, or at least epiphenomenalist dualism holds actually, or because physicalism does.

The second worry about my schmoperty move concerns the obscurity of the notion of something that seems to be a property but is not, i.e. the obscurity of the notion of schmoperty. In reply I would first ask: in virtue of what aren't notions like *schmass*, *schmelectron*, or *schmorality* considered obscure? Indeed, philosophers are happy to use such terms when it comes to naming some possibly instantiated alien property that otherwise we would prima facie be tempted to name by a rigid designator of an actually instantiated property. Well, it is pretty clear that we use these names in virtue of some resemblance between the property referred to by the rigid designator that designates an actually instantiated property and the possibly instantiated alien one. Now the objection against the notion of a schmoperty could then be that resemblance among any kind of things requires sameness of some of the things' *properties*, yet in a world inhabited by schmoperties there are, by stipulation, no properties, so we cannot make sense, or at least not in the standard way, of schmoperties.

However, it seems to me that what is important in cases when we introduce a new term, like 'twater', in order to accommodate the intuition that what is picked out by it is not water, but resembles it, or like 'schmelectron', in order to accommodate the intuition that what is picked out by it is not electron, but resembles it, is either a theoretical role, as in the latter case, or a folk functional role, as in the former, which is played by the possible entities under scrutiny. There is, however, no determinate commitment at this point to whether the fillers of these roles are fillers in virtue of shared properties, or in

virtue of shared tropes, or in virtue of belonging to the same set of particulars. What is important is that those *possibilia* do play the role. The same thing is therefore important in the case of calling something a ‘schmoperty’ in the context of the causally based essentialism about properties: it is a *possibile* that plays the/a<sup>14</sup> theoretical property role, where the theory in question is ontology. What the propounder of the causal essentialist approach to properties adds to this is that the notion of property role that she is working with is such that the relation between ‘property’ and ‘property role’ comes out as *a priori* and contingent, while that between ‘property’ and ‘having such and such conditional causal powers’ comes out as *a posteriori* and necessary. Of course, the notion of a theoretical role does not have to imply that of a causal role, the causal role being one of the ways to *realise* a theoretical role.

However, the objection can be pushed further. Even if we accept that the idea of schmoperties is intelligible by appeal to sameness of theoretical ontological role, the problem of intelligibility reappears at the level of terms for schmoperties as qualified by some adjective; for instance, we can only make sense of the 1-possibility of *zombies* in terms of instantiation of schmoperties rather than properties if schmoperties can be qualified by the adjective ‘physical’, or at least by ‘schmysical’, just as a requirement for making sense of some property being schmysical rather than physical is that the schmysical property be qualified as seeming (to be) physical (as, indeed, Stoljar’s object physicalism and Chalmers’ panprotopsychism appear to require). The above construal appealing to theoretical role can safely be extended to this level: what accounts for the

---

<sup>14</sup> Indeed, there are various alternative ontologies, and therefore theoretical property roles, as pointed out e.g. by Lewis, David K., *On the Plurality of Worlds*. Oxford: Basil Blackwell, 1986, p. 55. For more about the idea of property role see Oliver, ‘The metaphysics of properties’, 1996, pp. 14-20.

qualification 'schmysical' is, again, a theoretical physical role, which is not necessarily a causal role. Indeed, on one version of object physicalism/panprotopsyism, intrinsic categorical essences of properties described by physical theory – Stoljar's object physical properties, Chalmers' protophenomenal ones - ground (i.e. are necessary for) these latter properties' causal activity. If this is so, then it follows that we can make sense of a theoretical role of theory-physical properties without appeal to causation, and sameness of that theoretical role across possible worlds ensures the intelligibility of something that seems physical but is not, i.e. of something that is schmysical. I don't see then why not understand the required schmysicalness of schmoperties in exactly the same way. Schmysical schmoperties are then some items that satisfy the theoretical role of physical properties, but are neither physical, nor properties.<sup>15</sup>

The third objection concerns the fact that there are properties that we conceive of in no other way than as intrinsic, non-dispositional - contrary to how the causal account requires us to conceive of them. I take it that conceiving of them in this fashion does not entail that they don't have conditional powers. What is entailed is perhaps that we don't have the intuition of the essentiality of powers to these properties. But the Individuation Thesis and broad Kripkeanism accommodate this fact: the former does not require more than conditional powers as *a posteriori* necessarily connected to the properties they are

---

<sup>15</sup> Alternatively, one could speak in terms of *physical* schmoperties, if one subscribes to the Russellian line of thought according to which causation is not a notion that physics needs at all, so that Humean constant conjunction is enough for an entity to qualify as physical. Anyhow, this doesn't change the fact that if the causal essentialist metaphysics of properties is true, these are only schmoperties rather than properties, so the reply to the zombie argument still works.

conditional powers of<sup>16</sup>, while the latter explains why we have a counter-intuition regarding this necessary connection. If the objection worked it would work for ‘water is H<sub>2</sub>O’ as well: it would say that we do not have an intuition of essentiality of H<sub>2</sub>O to water since we conceive of water without reference to what actually composes it. The reply will be, of course, that conceiving of water that way explains the intuition against necessity or essentiality by way of a metaphysical possibility that is being misdescribed as the possibility of water not being H<sub>2</sub>O.

A fourth objection concerns what might be called “the substance of physicalism” as opposed to the particular argument discussed here, the one from the conceivability of zombies<sup>17</sup>. It says that even we accept the argument presented here, given that the causal theory of properties does not distinguish between the case in which zombie worlds are impossible because physicalism is true and the case in which they are impossible because necessitarian dualism (where mental and physical properties are distinct but necessarily connected) is true, the substance of the physicalism issue is not really addressed. In reply, I would point out, first, that our argument at least has weakened the case for the conceivability based arguments against physicalism, and, second, that we have found a new way in which physicalism could be true, which is useful in the actual context, where physicalists are mostly required to defend the coherence of their view rather than to offer positive arguments for it. Thirdly, there is indeed an important point made in this objection, namely, that for all the argument has proven, the substance of the issue of the truth or otherwise of physicalism either is and is going to remain a deep and inaccessible

---

<sup>16</sup> This is, broadly speaking, because the causal account is 1-conceivably false even if actually (and necessarily true). It also explains why we can stick, after running the argument, with the acceptance of the 1-conceivability of zombies.

<sup>17</sup> Thanks to an anonymous referee for pointing this out.



one, or, on the contrary, there is a flavour of unsubstantiveness to it. Both these further strategies are, in my view, perfectly viable to pursue, but I will leave them for another occasion.

## **5. Conclusion**

I have tried to show in this paper that even the most sophisticated analysis of the zombie argument, namely, the one based on two-dimensional semantics, fails to exhaust the logical space of possible replies. The reply I have proposed is based on the causal account of properties, which generates necessitarianism about laws and causation. A growing number of philosophers have recently expressed their sympathy for either the causal account or law and causation necessitarianism. There is, of course, disagreement about whether necessitarianism is warranted, but at least it is a coherent view, therefore, it has to be taken into account, with all its implications, even in the context of the mind-body problem. If I was right in this paper, the causal account of properties raises a genuine challenge to conceivability-based arguments against physicalism.