

Andrew Botterell

Conceiving What Is Not There

Abstract: *In this paper I argue that certain so-called conceivability arguments fail to show that a currently popular version of physicalism in the philosophy of mind is false. Concentrating on an argument due to David Chalmers, I first argue that Chalmers misrepresents the relation between conceivability and possibility. I then argue that the intuition behind the conceivability of so-called zombie worlds can be accounted for without having to suppose that such worlds are genuinely conceivable. I conclude with some general remarks about the nature of conceivability.*

I

Conceivability arguments have traditionally played a prominent role in philosophy of mind. Descartes' argument for the separation of mind and body is perhaps the most famous example of such an argument, but more recent discussions rely heavily on conceivability considerations as well.¹ Despite their prominence, however, it remains a matter of considerable debate whether conceivability arguments are capable of doing the work they are asked to do. Indeed, it seems to me that on reflection such arguments are a good deal less persuasive than many philosophers take them to be. My aim in this paper is thus to argue that conceivability arguments do not present problems for a certain attractive and widely endorsed version of physicalism in the philosophy of mind.

There are two issues that I want to address. First, there is a general question about the relation between conceivability and possibility. In brief: do conceivability considerations provide us with evidence about what is, and what is not, possible? And second, there is a more specific issue about whether so-called zombie worlds are conceivable. The paper is structured with these two issues in mind. I begin with a discussion of the version of physicalism that I am interested in defending. I then consider a conceivability argument due to David Chalmers for the conclusion that this version of physicalism is false. In the next part of the paper I criticize

Correspondence: Andrew Botterell, Department of Philosophy, Sonoma State University, Rohnert Park, CA 94928, USA. Email andrew.botterell@sonoma.edu

[1] Two influential examples are Kripke (1980) and Chalmers (1996).

Chalmers' argument against physicalism. I focus in particular on two aspects of his argument. I first argue that Chalmers misrepresents the relation between conceivability and possibility. I then argue that the intuition behind the conceivability of so-called zombie worlds can be accounted for without having to suppose that such worlds are genuinely conceivable. The upshot is that the purported conceivability of zombies does not present a problem for physicalism.

II

But first, some stage-setting. In *The Conscious Mind*, Chalmers argues for property dualism, the view that certain properties of conscious psychological experiences are not identical with any physical properties. Chalmers' main target is a certain view of the relation between physical nature and psychological nature, a view that I will call *a posteriori physicalism*.

A posteriori physicalism is the conjunction of two theses, one metaphysical, the other epistemological. I propose to take the following thesis as my target formulation of the metaphysical thesis of physicalism:

- (P) Physicalism is true of a world w iff any physical duplicate of w is a duplicate simpliciter of w .

This calls for explanation. Call a world w a *physical duplicate* of a world w^* if w is indistinguishable from w^* in all physical respects. Call a world w a *duplicate simpliciter* of a world w^* if w is indistinguishable from w^* in *all* respects. To say that physicalism is true of the actual world is thus to say that any world that is indistinguishable from the actual world in all physical respects is indistinguishable from the actual world in all respects.²

In addition, however, the claim that physicalism is true of the actual world also carries with it a commitment to the idea that the physical nature of the actual world *entails* the psychological nature of the actual world. To see what this means, suppose $\phi_{@}$ is the complex physical sentence that states the entire physical nature of the actual world. And suppose $\psi_{@}$ is the complex psychological sentence that states the entire psychological nature of the actual world. Then if physicalism is true of the actual world, it follows that the conditional $\lceil \phi_{@} \supset \psi_{@} \rceil$ is necessarily true: any world at which $\phi_{@}$ is true is a world at which $\psi_{@}$ is true. More generally, if physicalism is true of the actual world then for any true psychological sentence ψ there is a true physical sentence ϕ such that the conditional $\lceil \phi \supset \psi \rceil$ is necessarily true — that is, if physicalism is true of the actual world every psychological way the actual world is will be made true by some physical way the actual world is. For simplicity I will call this the *Physical Entailment Thesis*.

[2] For a fuller discussion of how best to formulate the metaphysical thesis of physicalism see among others Lewis (1983), Jackson (1998) and Byrne (1999). I will return to the issue of what makes two worlds physical duplicates below. Also, for the purposes of this paper I assume — somewhat controversially — that the metaphysical thesis of physicalism is coherent and not obviously false. For arguments that this assumption is mistaken, see Chomsky (1993) and Crane and Mellor (1990).

We are now in a position to state the thesis of a posteriori physicalism more precisely. A posteriori physicalism is the conjunction of two claims. First, the claim that physicalism is true of the actual world, and hence — by the Physical Entailment Thesis — that for every true psychological sentence ψ there is a true physical sentence ϕ such that the conditional $\lceil \phi \supset \psi \rceil$ is necessarily true. And second, the claim that the entailment of psychological nature by physical nature is only a posteriori. A posteriori physicalists therefore ‘accept that [psychological nature is] not necessitated a priori by physical [nature], but hold that [it is] necessitated a posteriori by physical [nature]’ (Chalmers 1999, p. 474). In short, a posteriori physicalism maintains that although there is a *necessary* entailment of psychological nature by physical nature, there is no *a priori* or *conceptual* connection between the two. This will become important later on.

A posteriori physicalism is an attractive thesis. Its metaphysical component makes sense of the idea that physical nature is in some sense basic; and its epistemological component accounts for the intuitively plausible idea that there is no a priori connection between our concepts of physical nature and our concepts of psychological nature. This aspect of a posteriori physicalism is controversial, however, and some philosophers have argued to the contrary that there is an a priori or conceptual link between our concepts of physical nature and our concepts of psychological nature.³ Although none of these arguments strikes me as persuasive, I will not address this issue here.

What I will focus on instead is an argument of Chalmers’ for the conclusion that a posteriori physicalism is false. The argument runs as follows:

*Chalmers’ Argument*⁴

- (P1) In the actual world, there are conscious experiences.
- (P2) There is a physical duplicate of the actual world in which the facts about consciousness in the actual world do not hold.
- (P3) Therefore, facts about consciousness are further facts about the actual world, over and above the physical facts.

So:

- (C) Physicalism is false of the actual world.

Given my assumptions concerning physicalism and the Physical Entailment Thesis, I agree that the argument is valid. What I wish to focus on, then, is the question whether the argument is sound. Central to this argument is the claim that there is a world that is physically identical with, but psychologically different from, the actual world. The questionable premise is therefore (P2). Chalmers’ reason for thinking that (P2) is true is that zombie worlds are conceivable, and hence metaphysically possible. I will argue to the contrary that Chalmers’ reasons for thinking that zombie worlds are metaphysically possible are flawed. In order to do this, however, I need to first talk about conceivability, possibility, and the relation between the two notions.

[3] See, for example, Lewis (1966; 1994), Armstrong (1968) and White (1986).

[4] See Chalmers (1996), p. 123.

III

It is natural to take the things that are claimed to be conceivable and possible to be propositions. Chalmers focuses on statements instead. For the purposes of this discussion, I will follow him in doing so. A statement, for Chalmers, is a sentence — or something very similar to a sentence — that has truth-conditions, or expresses a proposition. Chalmers distinguishes the conceivability of statements from the conceivability of propositions, or worlds. It is important for Chalmers, however, that he be able to move from the conceivability of a statement to the existence of a possible world corresponding to that statement. (‘Corresponding to’ is deliberately vague; I’ll attempt to make it more precise below.) To this end, Chalmers suggests the following: a statement E is conceivable (or conceivably true) if E is true in some conceivable, or logically possible, world.^{5,6} But Chalmers also thinks that if a statement E is conceivable, then E is possibly true, and hence, that there is a logically possible world in which E is true. For example, he remarks that ‘whenever it is conceivable that [a statement E is true], there will be a logically possible world in which [E is true]’ (Chalmers 1999, p. 482). Putting the two halves together: a statement E is conceivable if, and only if, there is a logically possible world in which E is true.

The interesting part of this biconditional for our purposes is the claim that a statement E is conceivable only if there is a possible world in which E is true. Call this (Con):

(Con) If a statement E is conceivable then E is possibly true.

I said above that it is important for Chalmers that he be able to move from the conceivability of a statement to the existence of a possible world ‘corresponding to’ that statement. (Con) allows him to do precisely that.

Let us return to Chalmers’ argument against a posteriori physicalism. Again, what is at issue is (P2), the claim that there is a world that is physically identical with, but psychologically different from, the actual world. Chalmers’ reason for thinking that (P2) is true is that zombies — ‘someone or something physically identical to me (or to any other conscious being), but lacking conscious experiences altogether’ (Chalmers, 1996, p. 95) — are conceivable, and hence, that zombie worlds are possible. Chalmers’ reason for finding zombie worlds conceivable is found in the following passage:

-
- [5] Although Chalmers talks of conceivable, or logically possible, worlds — and uses the two more or less interchangeably — this should not be taken to mean that he thinks that there are two distinct sets of worlds, the conceivable or logically possible ones, and the metaphysically possible ones, and that the latter are a subset of the former. The notion of a logically possible world functions as a primitive for Chalmers. As he says, ‘[a]s for the notion of a logically possible world, this is something of a primitive . . . we can intuitively think of a logically possible world as a world that God might have created’ (Chalmers, 1996, p. 66). However, since any world that God might have created is a possible world, it is hard to avoid the conclusion that by a logically possible world Chalmers simply means a possible world.
- [6] For ease of exposition I will sometimes say that a statement E is conceivable (or that E is possible). This is to be understood as short for: E is conceivably true (or E is possibly true).

the [conceivability] of zombies seems . . . obvious to me. A zombie is just something physically identical to me, but which has no conscious experience — all is dark inside. While this is probably empirically impossible, it certainly seems that a coherent situation is described; I can discern no contradiction in the description. In some ways an assertion of this logical possibility comes down to a brute intuition but . . . [a]lmost everybody, it seems to me, is capable of conceiving of this possibility. (Chalmers, 1996, p. 96)

In this passage, Chalmers suggests a test for determining whether a situation, or world, is conceivable. The test seems to be this: if no contradiction can be discerned in the description of a situation, then the situation so described is conceivable. So in particular, zombie worlds will be conceivable if there is no contradiction in the description of a situation, or world, which is physically indistinguishable from the actual world, but in which there are no conscious psychological states. Chalmers can thus be interpreted as offering the following argument for (P2):

Zombie Argument

- (Z1) It is conceivable that there is a world that is physically indistinguishable from the actual world but in which there is no consciousness, i.e., it is conceivable that there are zombie worlds.
- (Z2) If it is conceivable that there are zombie worlds, then it is possible that there are zombie worlds.

So:

- (ZC) It is possible that there are zombie worlds.

It should be granted that this argument has some intuitive force, especially against proponents of a posteriori physicalism. For since according to a posteriori physicalism there is no a priori connection between physical nature and psychological nature, the claim that zombie worlds are conceivable seems initially plausible. Moreover, it is also clear that the conclusion of the zombie argument presents problems for physicalism. For if zombie worlds are possible then physicalism is false of the actual world, as the existence of such worlds will entail the falsity of the claim that for every true psychological sentence ψ there is a true physical sentence ϕ such that the conditional $\lceil \phi \supset \psi \rceil$ is necessarily true. Nonetheless, despite the initial plausibility of Chalmers' zombie argument, I think that there are problems with it. I will argue, first, that Chalmers equivocates on the crucial notion of conceivability and that as a result, his argument is invalid. I will then argue, using some of Chalmers' own terminology, that zombies are not conceivable in the relevant sense, and hence, that (Z1) is false.

IV

I'll begin with (Con) — the claim that if a statement E is conceivable, then E is possible. This sort of view about the relation between conceivability and possibility is not without critics. Joseph Levine, for example, argues that 'what is

[conceivable] is an *epistemological* matter . . . It takes another argument to get from [an] *epistemological* possibility . . . to [a] metaphysical possibility' (Levine, 1993/1997, p. 544). Or again: 'suppose we reject the Cartesian model of epistemic access to metaphysical reality altogether. One's ideas can be as clear and distinct as you like, and nevertheless not correspond to what is in fact possible. The world is structured in a certain way, and there is no guarantee that our ideas will correspond appropriately' (Levine, 1993/1997, p. 544).

One way to interpret Levine's objection is that there is no route from conceivability to possibility, or in our terminology, that the conceivability of a statement *E* provides no evidence for *E*'s possible truth.⁷ On reflection, however, I think it is clear that this is not a tenable view. I say this because it seems clear that conceivability considerations *do* sometimes provide us with evidence about what is and what is not possible. For example, it seems conceivable to me that I could exist in the absence of Building 20, and it seems odd to deny that this provides me with evidence for thinking that it is possible that I could exist in the absence of Building 20. To deny that conceivability considerations ever provide us with evidence about what is possible seems implausible. What we want to know is why appeals to conceivability considerations are legitimate in some cases but not in others. To claim that they are never legitimate is to ignore this important question altogether.

Of course, it might be thought that we can provide arguments for the conclusion that the evidence provided by conceivability considerations is specious. Consider, for example, Descartes' argument for the separation of mind and body. Descartes was concerned to argue that persons are not physical objects. Descartes argued that 'from the fact that I have gained knowledge of my existence without noticing anything about my nature or essence except that I am a thinking thing, my essence consists solely in the fact that I am a thinking thing' (Descartes, 1993, p. 132). And since according to Descartes his body is not something whose essence consists solely in the fact that it is a thinking thing, he concluded that he is not identical with his body. The general principle at work in Descartes' argument seems to be this: if I do not notice that a property *P* is part of my essence, then *P* is not part of my essence.

Descartes' Principle

For all individuals *x*, and all properties *P*, if *x* does not notice that *P* is part of *x*'s essence, then *P* is not part of *x*'s essence.

A number of commentators have taken issue with Descartes' principle. For example, Arnauld asked how it is supposed to follow from the fact that I don't notice that a property *P* is part of my essence that *P* is not part of my essence.⁸ Or — in the particular case we are concerned with — how is it supposed to follow

[7] It's not clear to me that this is Levine's view, but it is nonetheless suggested by his remarks. Even if Levine doesn't hold this view, however, I think that it is a view to which many are sympathetic and so to that extent deserves consideration.

[8] See Yablo (1990) for a fuller discussion of Arnauld's objection, and Descartes' responses to it.

from the fact that I don't notice that my essence includes being physically embodied that being physically embodied is not part of my essence? Rejecting Descartes' principle would provide us with an attractive way to rebut his argument for the conclusion that persons are not physical objects.

As Stephen Yablo points out, however, rejecting Descartes' principle for Arnauldian reasons results in a highly implausible view about the evidential weight of conceivability intuitions generally. Again, it seems possible to me that I could have been born on a day other than the day I was in fact born on. If the line of argument just canvassed is any good, however, this reasoning is suspect. For how do I know, from the fact that I do not notice that the property of being born in Toronto is part of my essence, that that property is *not* part of my essence? Perhaps, despite my not noticing it, my essence is such that it includes within it that very property. Without something like Descartes' Principle in the background, its seeming to me that I could have been born in a different place gives me no reason to believe that I could have been born in a different place, and that seems highly implausible. The problem is perfectly general: rejecting Descartes' principle makes all modal intuitions suspect, and we have good reason to think that this cannot be right. Consequently, it seems to me to be a mistake to object to (Con) on the grounds that conceivability considerations never provide us with evidence about what is and what is not possible.

A more subtle objection to this sort of account of the relation between conceivability and possibility is due to James van Cleve (1983). Says van Cleve,

[w]hatever we mean by 'conceivable', we shall have to contend with the following objection. In 1742 Goldbach proposed to Euler that every even number greater than two is the sum of two primes — a conjecture that to this day has eluded both counterexample and proof. Now isn't it conceivable that Goldbach's conjecture is true? Most people to whom I have put his question say yes. But now consider the negation of Goldbach's conjecture: isn't it conceivable as well? Most people to whom I have put the question again say yes. Thus many people profess to find Goldbach's conjecture and its negation equally conceivable. But one of them is impossible! Since either Goldbach's conjecture or its negation must be true and since every proposition of mathematics is necessarily true if true at all, one proposition in this pair must be a necessary truth and the other an impossibility. It thus appears that at least one impossible proposition is conceivable (van Cleve, 1983, p. 36).

Van Cleve suggests that we distinguish two senses of conceivability — a positive and a negative sense — as follows:⁹

Strong Conceivability

A statement *E* is *strongly conceivable* for a subject *S* iff *S* sees that *E* is possibly true.

Weak Conceivability

A statement *E* is *weakly conceivable* for a subject *S* iff *S* does not see that *E* is necessarily false.

[9] Van Cleve's original definitions quantify over propositions rather than statements, but I have reformulated them to accord with Chalmers' usage.

Set to one side the question what it is to see that a statement is possible; I will return to it shortly.

With the distinction between Strong and Weak Conceivability in hand, van Cleve argues that Goldbach's conjecture is only weakly conceivable. On this view, when people claim to find Goldbach's conjecture (or its negation) conceivable, what they mean is that they do not see that Goldbach's conjecture is impossible (nor do they see that the negation of Goldbach's conjecture is impossible). However, since Goldbach's conjecture is necessarily true if it is true at all, we know that either Goldbach's conjecture or its negation is logically impossible. Thus, since at least one of the statements claimed to be weakly conceivable is necessarily false, we can conclude that it does not follow from a statement's being weakly conceivable that it is possibly true.

What about Strong Conceivability? If S sees that E is possibly true, does it follow that E is possibly true? The answer to this question depends on what it means to see that a statement is possible. It might be thought that if S sees that E is possibly true, then it immediately follows that E is possibly true. For if 'sees' is understood factively, a subject cannot see that a statement E is possibly true and yet it be the case that E is not possibly true. However, there is another sense of 'sees', a non-factive sense, according to which ' S sees that a statement E is possible' means 'it seems or appears to S that E is possible', from which it does not follow that E is possible.¹⁰ What is interesting about this is that, combining the positive and negative senses of 'is conceivable' suggested by van Cleve with the factive and non-factive senses of 'sees', we are left with *three* different senses of 'is conceivable'. For simplicity, let us stipulate that 'sees' is to be understood factively in the following definitions. Then we have the following three interpretations of 'is conceivable':

Factive Strong Conceivability

A statement E is *factively strongly conceivable* for a subject S iff S sees that E is possibly true.

Apparent Strong Conceivability

A statement E is *apparently strongly conceivable* for a subject S iff it appears to S that E is possibly true.

Weak Conceivability

A statement E is *weakly conceivable* for a subject S iff S does not see that E is necessarily false.¹¹

[10] In short, a verb V is factive if it follows from the truth of ' S Vs that P ' that the proposition that P is true. Thus, 'knows' is factive, since it follows from the truth of ' S knows that P ' that the proposition that P is true. 'Believes', on the other hand, is non-factive, since it does not follow from the truth of ' S believes that P ' that the proposition that P is true.

[11] In the case of weak conceivability the factive and non-factive interpretations of 'sees' collapse, since the claim that a subject S does not see that a statement E is necessarily false is equivalent to the claim that it does not appear to S that E is necessarily false.

Now, it should be uncontroversial that Weak Conceivability is a non-starter, since as we have seen it is consistent with Weak Conceivability that necessary falsehoods are conceivable.¹² So we are left with a choice between Factive Strong Conceivability and Apparent Strong Conceivability as definitions of 'is conceivable'. This gives us the following accounts of the relation between conceivability and possibility, which I will call (FSC) and (ASC), respectively:

- (FSC) If a statement *E* is factively strongly conceivable,
then *E* is possibly true.
- (ASC) If a statement *E* is apparently strongly conceivable,
this provides prima facie evidence for *E*'s being possibly true.

Moreover, each of (FSC) and (ASC) has something to recommend it. In particular, both account for the intuition that conceivability considerations provide us with evidence about what is and what is not possible.

Let us return to Chalmers' zombie argument. With our two different interpretations of the relation between conceivability and possibility in hand, we can offer the following two interpretations:

Zombie Argument (Factive)

- (Z1-F) It is factively strongly conceivable that there is a world which is physically indistinguishable from the actual world but in which there is no consciousness, i.e., it is factively strongly conceivable that there are zombie worlds.
- (Z2-F) If it is factively strongly conceivable that there are zombie worlds, then it is possible that there are zombie worlds.

So:

- (ZC-F) It is possible that there are zombie worlds.

and

Zombie Argument (Apparent)

- (Z1-A) It is apparently strongly conceivable that there is a world which is physically indistinguishable from the actual world but in which there is no consciousness, i.e., it is apparently strongly conceivable that there are zombie worlds.
- (Z2-A) If it is apparently strongly conceivable that there are zombie worlds, then it is prima facie possible that there are zombie worlds.

So:

- (ZC-A) It is prima facie possible that there are zombie worlds.

And our question is: which argument fares better?

[12] Indeed, Chalmers rejects Weak Conceivability for precisely this reason (see Chalmers, 1996, p. 67).

Now, I think it is clear that Zombie Argument (Factive) will not do, and for a simple reason. The problem is that although (Z2-F) bridges the gap between conceivability and possibility — owing, again, to the fact that ‘sees’ in (FSC) is being interpreted factively — there is good reason to think that (Z1-F) is false: it is simply not the case that it is factively strongly conceivable that there are zombie worlds. I realize that this amounts to an unargued assertion at the moment, and I will have much more to say about why I think that (Z1-F) is not forced on us below.

Zombie Argument (Apparent), on the other hand, is intuitively very plausible. At the very least, since it is not implausible to suppose that it is apparently strongly conceivable that there are zombie worlds, (Z1-A) seems reasonable enough. However, although (Z1-A) is plausible when ‘is conceivable’ is understood in the sense of ‘is apparently strongly conceivable’, the problem is that (Z2-A) is too weak to generate the conclusion that Chalmers is after, namely that it is possible that there are zombie worlds. For even if we grant that ‘there are zombie worlds’ is apparently strongly conceivable, this only provides us with *prima facie* evidence for the possibility that there are zombie worlds; something else is needed to bridge the gap between this claim and the claim that zombie worlds *are* possible.

The problem with Chalmers’ zombie argument, then, is that it equivocates on the phrase ‘is conceivable’. The reading of ‘is conceivable’ according to which the second premise of the argument is plausible — the Factive Strong Conceivability reading — is not the reading of ‘is conceivable’ according to which the first premise of the argument is plausible. I therefore conclude that Chalmers’ zombie argument is not compelling.

V

Thus far I have been concerned to argue that there is a gap between the claim that it is conceivable that there are zombie worlds and the claim that it is possible that there are zombie worlds. As against this, however, it might be objected that the claim that zombie worlds are conceivable but not metaphysically possible gives rise to problems for which there are no satisfactory answers.¹³ In particular, it might be objected that this view will lead to scepticism about modal intuitions generally. Let me address this issue briefly before continuing.

I have been arguing that the conceivable truth of a statement *E* is consistent with, but does not imply, the claim that *E* is possibly true. But then how are we to tell, for any conceivably true statement, whether there is a possible world corresponding to it? If we cannot be assured that there is a world corresponding to a statement which is conceivably true, then aren’t intuitions about modality and conceivability always going to be suspect? In order to determine whether this criticism is a good one we need an account of modal error. The best account of modal

[13] From this point on, ‘is conceivable’ is to be understood in the sense of ‘is apparently strongly conceivable’.

error that I know of is due to Stephen Yablo. Yablo remarks, concerning the evidence provided by the conceivable truth of a statement *E*, that

this *prima facie* evidence [for the possible truth of *E*] is defeated if there is not improbably a [statement *P*] such that (a) [*P*] is true, (b) if [*E*] is true, then [*P*] is impossible, and (c) [*E*] is conceivable because one was unaware of (a) and/or (b) (Yablo, 1992, p. 254, note 24).

In short: the *prima facie* evidence for the possible truth of ‘there are zombie worlds’ provided by the conceivability of ‘there are zombie worlds’ is defeated if there is not improbably a statement *P* such that (a) *P* is true, (b) if *P* is true, then it is not possible that ‘there are zombies-worlds’ is true, and (c) ‘there are zombie worlds’ is conceivable because one is unaware of (a) and/or (b). Is there such a *P*? Unsurprisingly, I think there is.

Suppose *P* = ‘physicalism is true of the actual world’. Then, I think, the *prima facie* evidence provided by the conceivability of ‘there are zombie worlds’ is plausibly defeated. For it is not improbable, first, that physicalism *is* true of the actual world and second, that the reason zombie worlds seem possible is that we are unaware that physicalism is true of the actual world.

This will no doubt prompt the objection that a similar argument can with equal effectiveness be run to show that the conceivability of zombie worlds defeats the *prima facie* evidence provided by the conceivable truth of physicalism. For suppose *P* = ‘there are zombie worlds’. Wouldn’t its truth defeat the *prima facie* evidence provided by the conceivability of ‘physicalism is true of the actual world’? I do not think that it would. For this objection ignores an important point, namely that there are reasons having nothing to do with its conceivability for supposing that physicalism is true of the actual world. For example, there are many a posteriori reasons, stemming largely from the successes of the natural sciences in explaining many natural phenomena, for thinking that physicalism is true. And since what is actually true is possibly true, these reasons constitute an argument for thinking that physicalism is possibly true.¹⁴

In short, I am not concerned to argue that dualism could not possibly be true; rather, I am merely concerned to argue that conceivability considerations of the sort appealed to by Chalmers do not show a posteriori physicalism to be false. I will return to this issue again below.

VI

Where are we? Section IV established that Chalmers’ zombie argument is invalid: on the most plausible interpretation of ‘is conceivable’, the possibility of zombie worlds does not follow from the conceivability of such worlds. I’m now going to switch gears and argue, using some of Chalmers’ own terminology, that zombie worlds are not in fact conceivable. This will suffice to show that Chalmers’ zombie argument, in addition to being invalid, is also unsound.

[14] These reasons are, of course, defeasible. For it also true that the natural sciences have yet to provide a satisfying account of the phenomenon of consciousness, and this surely provides some reason for thinking that physicalism is not true.

Chalmers' view is that 'we need a logically possible world for every ideally conceivable [statement]' (Chalmers, 1999, p. 482). To understand why he thinks this, we need to spend a bit of time discussing the two-dimensional semantic framework employed by Chalmers.¹⁵ This is because given the two-dimensional framework there are two senses in which a possible world might correspond to a given statement.

The two-dimensional framework takes as its starting point the idea that there are two ways to think about the meaning of linguistic expressions.¹⁶ According to two-dimensionalism, associated with every expression are two intensions, or meanings, a *primary intension* and a *secondary intension*. An intension is a function from possible worlds to referents. The primary intension of an expression *E* says what *E* would refer to in a possible world *w* had *w* turned out to be actual.¹⁷ The secondary intension of an expression *E*, on the other hand, says what *E* refers to in any possible world given what *E* actually refers to. Consequently, when Chalmers says that there must be a possible world corresponding to every conceivable statement, he might mean either of two things. On the one hand, he might mean that there must be a possible world corresponding to a statement conceived of, or evaluated, according to its primary intension; on the other hand, he might mean that there must be a possible world corresponding to a statement conceived of, or evaluated, according to its secondary intension.

Following Chalmers, let us say that a statement *E* is *1-conceivable* if it is conceivable according to its primary intension; and let us say that a statement *E* is *2-conceivable* if *E* is conceivable according to its secondary intension. Similarly, let us call a statement *1-possible* if it is true in a possible world when evaluated according to its primary intension; and let us call it *2-possible* if it is true in a possible world when evaluated according to its secondary intension. More generally, then, Chalmers' view seems to be the following: whenever two concepts are coreferential only a posteriori, this can only be because their primary intensions differ. But if their primary intensions differ, then there must be a possible world in which the concepts yield distinct extensions. So associated with every true a posteriori identity claim there is both a necessary and a contingent proposition. Again, using 'water' and 'H₂O' as our examples, we can say that 'water is not H₂O' is 1-conceivable and 1-possible, although it is neither 2-conceivable nor 2-possible. This is because it is plausible to suppose that 'water' and 'H₂O' have the same secondary intension, but different primary intensions. So although there are no worlds in the secondary intension of 'water' which are not in the secondary intension of 'H₂O', there are arguably worlds which are in the primary intension of 'water' but are not in the primary intension of 'H₂O'.

I said at the outset that I was interested in two issues. One was whether conceivability considerations provide us with evidence about what is, and what is not, possible. The other was whether zombie worlds are coherently conceivable.

[15] For discussion of the two-dimensional framework, see Stalnaker (1978), Davies and Humberstone (1980), Tichy (1983), Lewis (1994), Jackson (1998), Block and Stalnaker (1999) and Byrne (1999).

[16] Or concepts. I will use the two interchangeably in what follows.

[17] Following the terminology of Davies and Humberstone (1980).

However, it is one thing to argue that, in general, conceivability considerations provide us with defeasible evidence for the possible truth of a given statement. It is quite another thing to argue, with respect to the particular case of zombie worlds, that such worlds are conceivable. It is to this issue that I now turn.

In discussing the objection that ‘there are zombie worlds’ could be conceivable despite the fact that no zombie world corresponds to it, Chalmers remarks that

the only route available to an opponent here is to claim that in describing the zombie world as a zombie world, we are misapplying the concepts, and that in fact there is a conceptual contradiction lurking in the description. Perhaps if we thought about it clearly enough we would realize that by imagining a physically identical world we are thereby automatically imagining a world in which there is conscious experience. But then the burden on the opponent to give us some idea of where the contradiction might lie in the apparently quite coherent description. If no internal incoherence can be revealed, then there is a very strong case that the zombie world is logically possible (Chalmers, 1996, p. 99).

Is it possible to offer an argument for the conclusion that in describing a world as a ‘zombie world’ we are misdescribing it? Elaborating on an argument originally due to Richard Boyd (1980), I will argue that it is. I want to begin, however, by discussing a slightly different example.

In *Naming and Necessity* Saul Kripke considers a number of consequences of the thesis that proper names and natural kind terms are rigid designators.¹⁸ One consequence is that certain identities previously thought to be only contingently true turn out to be necessarily true. One such identity is (1):

- (1) heat = molecular K.E.

According to Kripke, because the terms flanking the identity sign in (1) are rigid designators, (1) is, if true, necessarily true. Nonetheless, there is an intuition that (1) is only contingently true.

As is well known, Kripke has an explanation of why (1) appears contingent. According to Kripke, before we knew what heat was we knew which sensations it gave rise to. In particular, heat is something we identified by virtue of the fact that it gives rise to something that we call ‘the sensation of heat’. This is not to say that ‘heat’ *means* ‘whatever gives rise to the sensation of heat’. Rather, the sensation of heat is the property by which we initially identified the phenomenon of heat. Thus, when we discovered that heat = molecular K.E., we

discovered an identification which gives us an essential property of this phenomenon. We . . . discovered a phenomenon which in all possible worlds will be molecular motion — which could not have failed to be molecular motion, because that’s what the phenomenon [of heat] is. On the other hand, the property by which we identify it originally, that of producing such and such a sensation in us, is not a necessary property [of heat] but a contingent one (Kripke, 1980, p. 133).

[18] An expression α is said to rigidly designate an object x just in case α refers to x in every world in which x exists, and does not refer to something other than x in worlds in which x does not exist (see Kripke, 1980, pp. 48–9).

According to Kripke, it is this contingent property of heat that accounts for our intuition that heat might not have been molecular K.E. When (1) is imagined to be contingent — and so false — what is being imagined is a world in which something other than molecular K.E. gives rise to sensations of heat. But this is not a counter-example to (1), since it is not a world in which heat = molecular K.E. Rather, it is a world in which something else has the property — contingently possessed by molecular K.E. in this world — of giving rise to sensations of heat. But since ‘heat’ does not mean ‘that which gives rise to sensations of heat’ such a world is compatible with the truth of (1).

What Kripke offers is a strategy for redescribing the conceivable falsity of certain necessary truths in a way that does not reflect badly on the modal status of those truths. To recapitulate, whenever we have a necessary a posteriori truth, there is always going to be an intuition that the truth is only contingent — that is, because the truth is only a posteriori it will seem conceivably false. So if we want to insist that a certain statement is both a posteriori and necessarily true, we need to redescribe it in a way that accounts for the appearance of contingency but does not reflect badly on its modal status. In the case of (1), we can do this by finding a contingent property of heat via which reference to it was initially fixed, and arguing that what we are imagining when we are imagining that (1) is false is that something other than molecular K.E. has that contingent property.

What does Kripke’s redescription strategy have to do with the two-dimensional framework employed by Chalmers? Simply put, Kripke’s contingent reference-fixing properties are Chalmers’ primary intensions.¹⁹ In general, we can recover the primary intension of an expression containing rigid designators by replacing one or more of the rigid designators with the descriptions by which the reference of the rigid designators was initially fixed. For example, the primary intension of ‘heat’ is a function which picks out the thing around here that gives rise to sensations of heat. This is because that description arguably represents the way in which we think about the reference of ‘heat’. Hence, when evaluated according to the primary intension of ‘heat’, ‘heat is molecular K.E.’ becomes ‘the thing around here that gives rise to sensations of heat is molecular K.E.’. The differences between the two-dimensional view advocated by Chalmers and Kripke’s reference-fixing proposal therefore seem to be a matter of terminology.²⁰

Now, what goes for necessary a posteriori identities also goes, *mutatis mutandis*, for necessary a posteriori conditionals of the sort to which I have argued physicalism is committed. For ease of exposition, let’s focus on a simple conditional linking physical nature and psychological nature:

- (2) If x is in a brain state of type C then x is conscious.²¹

Following Shoemaker (1981), let us assume that being a brain state of type C is a *total realization* of the property of being conscious. That is, let us assume that

[19] For a nice discussion and comparison of Chalmers and Kripke on this issue, see Yablo (1999; 2000).

[20] Witness Chalmers: a ‘difference in primary intensions will correspond to a difference in reference-fixing properties’ (Chalmers, 1996, p. 373, note 15).

[21] This is obviously *very* much simplified; however, it will suffice to make the point I want to make.

brain states of type *C* are unconditionally sufficient for being conscious, so that there is no world in which a person is in a brain state of type *C* but is not conscious (owing to the fact that there is some other property that that person lacks, the instantiation of which is required for the instantiation of the property of being conscious). Hence, on the assumption that physicalism is true of the actual world it follows that (2) is, if true, necessarily true: any world in which the antecedent of (2) is true is a world in which the consequent of (2) is true. However, on the assumption that a posteriori physicalism is true it will seem possible that there could be worlds in which there are brain states of type *C* but no consciousness, that is, it will be conceivable that (2) is only contingently true.

Chalmers argues that neither two-dimensionalism nor Kripke's redescription strategy can help the friend of a posteriori physicalism here, but it is unclear to me that he is right to say this. To see why, let us grant, with Chalmers, that the primary and secondary intensions of the concept of being conscious coincide, that

[w]hat it takes for a state to be a conscious experience in the actual world is for it to have a phenomenal feel . . . if something feels like conscious experience, even in some counterfactual world, it *is* a conscious experience. All it means to be a conscious experience, in any possible world, is to have a certain feel (Chalmers, 1996, p. 133).

Let us focus instead on the concept of being in a brain state of type *C*.

According to Chalmers' own principles, the concept of being in a brain state of type *C* is a paradigm example of a physical concept. Because it is a physical concept, however, it is very plausible to suppose that it — like the concept of being H₂O, or the concept of having molecular K.E. — has two distinct intensions. Its primary intension reflects the way in which reference to brain states of type *C* was initially fixed; presumably, this was via a cerebroscope or some similar measuring device. The secondary intension of the concept of being in a brain state of type *C*, on the other hand, picks out brain states of type *C* in virtue of certain of their essential properties. Our first observation, then, is that the primary and secondary intensions of the concept of being in a brain state of type *C* are distinct.

It should also be granted, however, that the primary intension of the concept of consciousness is distinct from the primary intension of the concept of being in a brain state of type *C*. For on the assumption that a posteriori physicalism is true it follows that there is no logical connection between physical nature and psychological nature. So in particular, on the assumption that a posteriori physicalism is true it follows that there is no logical connection between the concept of being conscious and the concept of being in a brain state of type *C*. This is our second observation.

I have argued that the primary intension of the concept of being in a brain state of type *C* is distinct from its secondary intension, and also from the primary intension of the concept of being conscious. It is therefore open to physicalists to argue as follows: a world that appears to contain brain states of type *C* but no consciousness is really a world in which the primary intension of the concept of being in a brain state of type *C* is satisfied but in which neither the primary nor the secondary intension of the concept of consciousness is satisfied. That is, it is open to

physicalists to claim that what is being imagined to be true when it is imagined that zombie worlds are possible is not (3):

- (3) x is in a brain state of type C , and x is not conscious.

but is instead (4):

- (4) x is in a state that gives rise to readings of type R on a cerebroscope, and x is not conscious.

(4) is the result of reinterpreting (2) according to the primary intension of the concept of being in a brain state of type C . But (4)'s being true does not constitute a counterexample to the claim that physicalism is true of the actual world, since a world of which (4) is true need not be a world in which there are brain states of type C in the absence of consciousness. In short, Chalmers' argument for the conceivability of zombie worlds is plausible only because he does not take seriously the redescription strategy according to which the primary intension of the concept of being in a brain state of type C is at issue. Chalmers argues that since the primary intension and the secondary intension of the concept of being conscious coincide, there is no way to redescribe (2) in a way that does not reflect badly on the claim that it is necessarily true. But if we simply ignore the concept of being conscious and focus instead on the concept of being in a brain state of type C , the redescription strategy is applicable. Consequently, since it is always open to physicalists to redescribe worlds which seem physically indistinguishable from the actual world as worlds which only appear to be physically indistinguishable from the actual world, Chalmers' argument for the conceivability of zombie worlds loses its force.

VII

There are two obvious responses to this argument. First, somebody might argue that it is an essential property of brain states of type C — or whatever physical property consciousness is said to be identical with, or supervene on — that they give rise to readings of type R on cerebroscopes. This objection is not to the point, however. For the issue is not whether brain states of type C necessitate readings of type R on cerebroscopes, or what have you. Rather, the issue is whether there could be readings of type R in the absence of brain states of type C . So long as this possibility is acknowledged the redescription argument remains unaffected.

Second, it might be objected that the claim that putative zombie worlds are physically distinguishable from the actual world is implausible. Chalmers himself presses this objection.²² I focused on a simple conditional linking physical nature and psychological nature. But suppose we focus instead on a more complicated conditional, for example, $\lceil \phi_{@} \supset \psi_{@} \rceil$ (where again, $\phi_{@}$ is the complex physical sentence that states the entire physical nature of the actual world, and $\psi_{@}$ is the complex psychological sentence that states the entire psychological nature

[22] See Chalmers (1996, pp. 134–6).

of the actual world). Then it might be argued that the redescription strategy described above is not applicable.

This objection has two parts. The first part alleges that, contrary to what I have argued, there is not a difference between the primary and the secondary intensions of $\phi_{@}$. The second part alleges that if the primary and the secondary intensions of $\phi_{@}$ do diverge, then it follows that at least some aspects of the physical nature of the actual world are essentially unknowable.

This is a powerful objection. What it is not, however, is convincing. As for the first part of the objection, since $\phi_{@}$ is a physical sentence, it is very plausible to suppose that it has two distinct intensions. For as I mentioned above, it is very plausible to suppose that physical concepts have their references fixed in whatever way that the concepts of scientific theories have their references fixed. If this is the case, then we have our distinction between primary and secondary intensions.

The second objection is harder to rebut. I have been arguing that we can account for the apparent conceivability of zombie worlds without having to admit that such worlds are genuinely conceivable. And I suggested that we can do so by redescribing the relevant necessary conditionals using the primary intensions of the physical concepts involved. The problem with this strategy, however, is that it seems to have the unwanted consequence that, for some physical concepts, we can only know about the physical nature of the actual world according to the primary intensions of those concepts. But if essential properties are tied to secondary intensions, this seems to preclude our having complete knowledge of the physical nature of the actual world. As Chalmers puts it,

it might be that for something qualify as an electron in a counterfactual world, it is not sufficient that it be causally related to other physical entities in the way that an electron [actually] is. Some hidden essence of electronhood might also be required. . . . The essential nature of electrons . . . would then be hidden to physical theory, which characterizes electrons . . . only extrinsically (Chalmers, 1996, p. 135).

It is important to recognize, however, that odd as it may seem, this sort of view is not without precedent. For there is a tradition in philosophy of thinking of the concepts and properties of physical theory as being dispositional in nature.²³ Classic examples of this sort of view can be found in Russell's (1912) presentation of the nature of matter, in his (1927) discussion of the nature of physical theory, and more recently, in the approach to the definition of theoretical terms due in part to Lewis (1970). If, however, this view is conjoined with the further view (i), that physical theory finds *only* dispositional properties, and (ii), that dispositional

[23] What is a dispositional property? Consider a simple example: the property of being fragile. Fragility is plausibly a dispositional property, since something is fragile if it is disposed to break when dropped. Or consider the property of being water-soluble: something is water-soluble if it is disposed to dissolve when immersed in water. More formally, we can say that an object x has a dispositional property D just in case x is such that, if x were in appropriate circumstances C , x would F , where F is what might be called the *manifestation* of D . It is a matter of considerable debate exactly which properties dispositions ought to be identified with, and there are reasons for thinking that the simple counterfactual account of dispositions just given requires revision. Nevertheless, this simple account will do for present purposes.

properties require non-dispositional properties as categorical bases, then we have a view according to which there are physical properties which lie outside the domain of physical theory.²⁴ Let us call this the *dual view*, since it amounts to the claim that there are two sorts of physical properties, those which are describable by physical theory — i.e., the dispositional properties picked out by the predicates of physical theory — and those which are hidden from physical theory — i.e., the non-dispositional, categorical bases, of those dispositional properties.²⁵

To be fair, Chalmers is not unaware of the possibility of this sort of view, and calls such hidden physical properties *protophenomenal properties*. His objection, however, is that this dual view collapses into a form of dualism which is no different from the view that he is arguing for. As he says, commenting on the dual view,

[a]fter ensuring that a world is identical to ours from the standpoint of our physical theories, God has to expend further effort to make that world identical to ours across the board. The dualism of ‘physical’ and ‘nonphysical’ properties is replaced on this view by a dualism of ‘accessible’ and ‘hidden’ physical properties, but the essential point remains (Chalmers, 1996, p. 136).

It pays, however, to ask what the essential point in question is. If the essential point is that the dual view commits us to a dualism of physical properties, then the proper response to this observation on the part of the physicalist should be to grant it. After all, it is unclear how it can be an objection to physicalism that certain properties, although physical, are nonetheless epistemologically hidden. If, on the other hand, the essential point is that *any sort* of dualism of properties shows that physicalism is false, then I think that the physicalist has reason to object. For the dual view will be equivalent to Chalmers’ dualism only if we make the further assumption that two worlds which are indistinguishable from the standpoint of physical theory are physical duplicates. But since this is precisely what is at issue, Chalmers cannot simply assume that it is true. What I am suggesting, in short, is that two worlds which are duplicates from the standpoint of physical theory needn’t be physical duplicates, since there may very well be additional physical respects in which such worlds differ. After all, there is certainly no a priori reason to suppose that physical theory must be capable of explaining *all* physical features of the world, and we have seen that there is an argument that physical theory — due to the fact that it invokes only dispositional properties and concepts — cannot do so. Thus, it seems to me that the idea that there are aspects of physical nature that are not accessible to physical theory is defensible and hence, that

[24] See Blackburn (1992) for an articulation of the view that it is ‘dispositions all the way down’. Armstrong (1968), while endorsing the idea that physical theory is concerned with dispositional properties, argues that physical theory does not tell us about the categorical bases of such properties. Smith and Stoljar (1998) and Stoljar (forthcoming) contain interesting discussions and defences of this sort of view. Jackson (1998) briefly, and inconclusively, discusses this view under the name *Kantian Physicalism*.

[25] One could resist this line of reasoning if one thought, along with Sydney Shoemaker (1980; 1998) that a property’s causal profile was essential to it, and thus, that any two properties with identical causal profiles are identical. Shoemaker’s proposal is interesting, but it too is saddled with mystery, since it makes it unclear what the things are that have causal profiles.

the dual view described above does not collapse into the sort of dualism that Chalmers is concerned to argue for.

Nonetheless, even if we can respond to Chalmers' worry, the dual view response is open to another objection. I have been arguing that zombie worlds might seem conceivable because we mistakenly think that certain worlds which are physically distinguishable from the actual world are in fact physically indistinguishable from the actual world. But on reflection it might seem that this is a non-starter. For if it is conceivable that there is a world which is physically just like the actual world in all accessible physical respects, and yet is mentally different, then surely it is also conceivable that there is a world which is physically just like the actual world in both accessible *and* hidden physical respects, and yet is mentally different. In other words, doesn't the appeal to hidden physical properties simply drop out?

In order to respond to this worry, we need to reflect again on the nature of conceivability.²⁶ It seems clear that in order to determine whether a statement is conceivable, we must be in a position to understand the concepts involved in the statement. For example, in order to determine whether it is conceivable that something could be a flurg without being a fish, we would need to understand the concept 'flurg', and know what property that concept picks out. If we are not in a position to understand the concept 'flurg', then we are not in a position to say whether or not it is conceivable that there could be a flurg which is not a fish. Similarly, in order to determine whether it is conceivable that there is a world which is physically just like the actual world in both accessible *and* hidden physical respects, and yet is mentally different, we would have to be in a position to understand the concepts which pick out the non-dispositional, hidden physical properties.

Now, I grant that we have concepts which pick out accessible physical properties; those are just the dispositional concepts of physical science. However, since the hidden physical properties are not expressible by the concepts of physical theory, and since those would appear to be the only concepts we are in a position to use to investigate the physical nature of the actual world, it is plausible that we *lack* the concepts required to pick out those hidden physical properties. And what follows from this is that it becomes unclear whether we can have any modal intuitions about the hidden physical properties. In particular, it becomes unclear whether we can have any intuitions about whether a world which is physically just like the actual world in both accessible *and* hidden physical respects could be mentally different.

Of course, somebody might object that for all I have argued, the concepts of non-dispositional physical properties might nonetheless be concepts that we are in a position to understand. But this is not to the point. For the issue is whether somebody who espouses the dual view is committed to thinking that we *must* be in a position to understand the concepts of non-dispositional physical properties. So far we have seen no reason to suppose that this must be the case.

[26] I am indebted here to discussions with Daniel Stoljar, and to Stoljar (forthcoming).

In sum, it seems to me that there is a way to account for the appearance of conceivability of the statement ‘there are zombie worlds’ without having to acknowledge the genuine conceivability of zombie worlds. I have tried to argue that it is possible to do so if we employ a version of Kripke’s redescription strategy. This strategy was initially introduced in the context of necessary identity claims, but I think that it can be deployed with equal effectiveness in the case of entailment claims of the sort to which I have argued physicalism is committed. To be sure, this requires that we recognize a distinction between two sorts of physical properties, and that we take seriously the possibility that the complete physical nature of the actual world will be hidden from view, but I see no reason to rule this possibility out a priori.

VIII

Let me conclude with some more general remarks about conceivability. I have been arguing that, on reflection, zombie worlds are not genuinely conceivable. Why then do so many claim to find such worlds conceivable? One possibility is that they have failed to appreciate the redescription strategy outlined above. But another, simpler, reason is plausibly the following. To conceive of a world which is physically identical with the actual world might seem easy enough to do. And to conceive of a creature which lacks consciousness wouldn’t appear to be difficult either: simply imagine a creature for whom ‘all is dark inside’. So to conceive of a zombie world I need only conceive of a world which is physically identical with the actual world, and which contains creatures for which all is dark inside. Does it follow that I have thereby succeeded in conceiving of a zombie world? I have certainly succeeded in imagining what a world would have to be like in order for it to be a zombie world; but to do that, I maintain, is not to conceive of a zombie world. For it does not follow from the fact that an *A*-world is conceivable, and that a *B*-world is conceivable, that an (*A*&*B*)-world is conceivable. For example, I can conceive of a world in which all objects are round; and I can conceive of a world in which all objects are square; but I cannot conceive of a world in which all objects are both round and square.

What makes it particularly difficult to determine whether a statement is genuinely conceivable is that it is an open question which background conditions should be held fixed.²⁷ Suppose somebody claims that it is conceivable that the moon should be made of green cheese. Is this really conceivable? Holding everything *except for the fact that the moon is made of green cheese* fixed, it might be thought that it is. But the fact is that we cannot hold everything fixed if our attempt to conceive that the moon should be made of green cheese is to be a serious one. In this I agree with Peter van Inwagen. Van Inwagen remarks that

anyone who thinks he can imagine that the moon is made of [green] cheese has a very sluggish imagination: the active imagination demands a pasture for the antecedently necessary thousands of thousands of millions of cows, demands a way to preserve a piece of cheese in broiling heat, freezing cold, and vacuum for thousands of millions

[27] What follows owes much to Sneddon (1972).

of years, demands some off-stage *machina* to protect a piece of cheese thousands of miles across from gravitational compression into non-cheese But any *serious* attempt to imagine the moon being made of green cheese must, like the unimaginable object itself, soon collapse under its own weight (van Inwagen, 1979, pp. 671–12).

The example is somewhat fanciful, but van Inwagen's point is not. Depending on what background assumptions are held fixed, some statements will seem initially conceivable. But as soon as we begin to consider what follows from the supposed conceivability of a statement *E*, it will often become unclear that what we are being asked to conceive of is genuinely conceivable. The case of the moon being made of green cheese is one such example; the case of zombie worlds is plausibly another. So while somebody might object that an inability to conceive of or imagine zombie worlds indicates a defect in imagination, it seems to me that, on the contrary, such an inability indicates instead an imagination which is over-active.

IX

To conclude, I have tried to show that Chalmers' argument from the conceivability of zombie worlds to the conclusion that physicalism is false fails. I argued, first, that Chalmers' zombie argument equivocates on the crucial notion of conceivability. I then argued that there is a way to account for the apparent conceivability of zombie worlds without having to embrace the claim that zombie worlds are genuinely conceivable. Finally, I discussed some general worries about the notion of conceivability itself. As I said at the outset, it is a matter of considerable debate whether conceivability arguments in the philosophy of mind are capable of doing the work they are asked to do. The considerations of the present paper suggest that they are not.²⁸

References

- Armstrong, D. (1968), *A Materialist Theory of the Mind* (London: Routledge).
 Blackburn, S. (1992), 'Filling in space', *Analysis*, Supp. Vol., pp. 60–5. Reprinted in his *Essays in Quasi-Realism* (Oxford: Oxford University Press, 1993).
 Block, N., Flanagan, O. and Güzeldere, G. (ed. 1997), *The Nature of Consciousness: Philosophical Debates* (Cambridge, MA: MIT Press).
 Block, N. and Stalnaker, R. (1999), 'Conceptual analysis, dualism, and the explanatory gap', *The Philosophical Review*, **108**, pp. 1–46.
 Boyd, R. (1980), 'Materialism without reductionism: What physicalism does not entail', in *Readings in the Philosophy of Psychology, Vol. 1.*, ed. N. Block (Cambridge, MA: Harvard University Press).
 Byrne, A. (1999), 'Cosmic hermeneutics', *Philosophical Perspectives*, **13**, pp. 347–83.
 Chalmers, D.J. (1996), *The Conscious Mind* (New York: Oxford University Press).
 Chalmers, D.J. (1999), 'Materialism and the metaphysics of modality', *Philosophy and Phenomenological Research*, **59**, pp. 473–96.
 Chomsky, N. (1993), *Language and Thought* (London: Moyer Bell).
 Crane, T. and Mellor, D.H. (1990), 'There is no question of physicalism', *Mind*, **99**, pp. 185–206.
 Davies, M. and Humberstone, I.L. (1980), 'Two notions of necessity', *Philosophical Studies*, **38**, pp. 1–30.

[28] I am very much indebted to Alex Byrne, David Chalmers, Robert Stalnaker, Daniel Stoljar, and Judith Jarvis Thomson for helpful criticism and advice. I would also like to thank the Editor-in-chief of this journal, and two anonymous referees, for comments that improved the paper.

- Descartes, R. (1993), *Meditations on First Philosophy*, in *Introduction to Philosophy*, ed. J. Perry and M. Bratman (Oxford: Oxford University Press, 1991).
- Jackson, F. (1998), *From Metaphysics to Ethics* (Oxford: Oxford University Press).
- Kripke, S. (1980), *Naming and Necessity* (Cambridge, MA: Harvard University Press).
- Levine, J. (1993/1997), 'On leaving out what it's like', in Block *et al.* (1997).
- Lewis, D. (1966), 'An argument for the identity theory', *Journal of Philosophy*, **63**, pp. 17–25. Reprinted in his *Philosophical Papers, Vol. 1.* (Oxford: Oxford University Press, 1983).
- Lewis, D. (1970), 'How to define theoretical terms', *Journal of Philosophy*, **67**, pp. 427–66.
- Lewis, D. (1983), 'New work for a theory of universals', *Australasian Journal of Philosophy*, **62**, pp. 343–77.
- Lewis, D. (1994), 'Reduction of mind', in *A Companion to the Philosophy of Mind*, ed. S. Guttenplan (Oxford: Blackwell).
- Russell, B. (1912), *The Problems of Philosophy* (Indianapolis: Hackett).
- Russell, B. (1927), *The Analysis of Matter* (London: Kegan Paul).
- Shoemaker, S. (1998), 'Causal and metaphysical necessity', *Pacific Philosophical Quarterly*, **79**, pp. 59–77.
- Smith, M. and Stoljar, D. (1998), 'Global response-dependence and noumenal realism', *The Monist*, **81**, pp. 85–111.
- Sneddon, G. (1972), 'Logical possibility', *Mind*, **81**, pp. 481–94.
- Shoemaker, S. (1980), 'Causality and properties', in *Time and the Cause*, ed. P. Van Inwagen (Dordrecht: D. Reidel Publishing Co.). Reprinted in his *Identity, Cause, and Mind* (Cambridge: Cambridge University Press, 1984).
- Shoemaker, S. (1981), 'Some varieties of functionalism', *Philosophical Topics*, **12**, pp. 83–118. Reprinted in his *Identity, Cause, and Mind* (Cambridge: Cambridge University Press, 1984).
- Stalnaker, R. (1978), 'Assertion', in *Syntax and Semantics, Vol. 9.*, ed. P. Cole (New York: New York Academic Press). Reprinted in his *Context and Content* (Oxford: Oxford University Press, 1999).
- Stoljar, D. (Forthcoming), 'Two conceptions of the physical', *Philosophy and Phenomenological Research*.
- Tichy, P. (1983), 'Kripke on necessity a posteriori', *Philosophical Studies*, **43**, pp. 225–41.
- van Cleve, J. (1983), 'Conceivability and the Cartesian argument for dualism', *Pacific Philosophical Quarterly*, **64**, pp. 35–45.
- van Inwagen, P. (1979), 'Review of *The Coherence of Theism*, by R. Swinburne', *The Philosophical Review*, **87**, pp. 668–70.
- White, S. (1986), 'Curse of the qualia', *Synthese*, **68**, pp. 333–68. Reprinted in *The Unity of the Self* (Cambridge: MIT Press, 1991).
- Yablo, S. (1990), 'The real distinction between mind and body', *Canadian Journal of Philosophy*, Supp. Vol. **16**, pp. 149–201.
- Yablo, S. (1992), 'Mental causation', *Philosophical Review*, **101**, pp. 245–80.
- Yablo, S. (1993), 'Is conceivability a guide to possibility?', *Philosophy and Phenomenological Research*, **53**, pp. 1–42.
- Yablo, S. (1999), 'Concepts and consciousness', *Philosophy and Phenomenological Research*, **59**, pp. 455–63.
- Yablo, S. (2000), 'Textbook Kripkeanism and the open texture of language', *Pacific Philosophical Quarterly*, **81**, pp. 98–122.

Paper submitted May 2000