

A Priori Conditionals and the Conceivability of Zombies

Raamy Majeed
The University of Otago
Philosophical Papers (forthcoming)

Abstract: The conditional analysis of phenomenal concepts, as proposed independently by Hawthorne (2002), Stalnaker (2002) and Braddon-Mitchell (2003), is used to undermine Chalmers's (1996) conceivability argument in a way that is compatible with our having the zombie intuition. The aim of this paper is twofold: (1) To remove current misconceptions concerning how the analysis is to be applied. It will be explained how there are two distinct ways the analysis can be used to undermine the conceivability argument. (2) To employ this exposition to defend the analysis from objections leveled against it by Chalmers (2005) and Alter (2007).

1. Introduction

According to the conditional analysis, what phenomenal concepts refer to depends on what exists in the actual world: if there actually are nonphysical states of the relevant type, our phenomenal concepts refer to such states, and if there are no such states, these concepts refer to physical states. Advocates of this analysis, such as Hawthorne (2002), Stalnaker (2002) and Braddon-Mitchell (2003), propose that it is advantageous in that it enables us to undermine Chalmers's (1996) conceivability argument in a way that is compatible with our having the zombie intuition.

Despite some initial interest, this response has, largely, fallen under the radar. The first reason for this is that there is significant confusion over how exactly the analysis is supposed to undermine the conceivability argument. For example, there is dispute over which premise of the argument the analysis allows us to contest¹, and confusion over which of its features are required to undermine the argument.² This is

¹ E.g. Alter (2007) and Chalmers (2005) seem to suppose that the proponents of the analysis deny the first premise, whereas Haukioja (2008) claims that they deny the second.

² I sometimes hear skepticism about the analysis based on the misconception that it requires a commitment to 2-dimensional semantics.

not surprising given that the proponents of the analysis differ over the details of the analysis and its employment in undermining the conceivability argument.³

The second is that there is a widely held but unfortunate misconception that the two objections leveled against the analysis by Chalmers (2005) and Alter (2007) are successful. Haukioja (2008) offers us a plausible defense of the analysis from these objections, which has not yet been rebuked. My view is that while Haukioja doesn't succeed in undermining these objections, neither objection is devastating.⁴

This paper aims to rectify these two problems. To that end, I shall attempt to remove the aforementioned confusion by providing an overview of the two distinct ways the analysis allows us to have the zombie intuition, as well as the two subsequent ways it enables us to undermine the conceivability argument. I shall then use this exposition to illustrate how neither objection to the analysis proves devastating.

The paper's structure is as follows. I shall outline the conceivability argument (§2) and the conditional analysis (§3). Then I shall explain how the analysis enables us to have the zombie intuition (§4), and undermine the conceivability argument (§5). Finally, I shall defend the analysis from the two main objections (§6).

2. The Conceivability Argument

The conceivability argument, as newly formulated by Chalmers (2005), is as follows:

- Z1) $P \ \& \ \sim Q$ is conceivable.
- Z2) If $P \ \& \ \sim Q$ is conceivable, it is metaphysically possible.

³ It isn't even clear whether Stalnaker's response involves a conditional analysis. Stalnaker himself doesn't explicitly claim that our phenomenal concepts are conditional. What he claims is that we are ignorant of a totality fact to the effect that the physical truths about our phenomenal experiences exhaust all truths about these experiences. If we were thus ignorant, the correct way to analyze our phenomenal concepts, arguably, is the conditional one.

⁴ There is a third objection, due to Alter, which contests Hawthorne's oracle argument that bolsters the second conditional. Since this objection doesn't contest the analysis itself or its application to the conceivability argument, it isn't devastating. Hence, I shall not discuss it here. See Alter (2007: 243) for the objection and Haukioja (2008:149) for a defense.

Z3) If $P \& \sim Q$ is metaphysically possible, physicalism is false.

Z4) Therefore, physicalism is false. (Chalmers 2005: 142)

Let P be the sum of all the physical truths at the actual world, and Q be some arbitrary phenomenal truth at the actual world. $P \& \sim Q$, then, is a world where all the physical truths at our world obtain, but where some arbitrary phenomenal truth about our world doesn't. In other words, such a world is a zombie world: a world that is exactly like ours physically but which is distinct with regards to phenomenal consciousness or qualia.⁵

A caveat. Zombie worlds, as Chalmers (1996) originally understands them, are worlds which are exactly like ours physically but which lack any phenomenal consciousness whatsoever. Subsequently, 'Q' in the original conceivability argument is read as the sum of all the phenomenal truths at our world. There are two reasons to prefer the new reading. First, (Z1), read in the old way, makes a much stronger claim, which is highly controversial.⁶ Second, as we shall now see, all we need to undermine physicalism is the new weaker claim.

Premise (Z1) states that zombie worlds are conceivable. In this context, to say that a given world (or epistemic scenario) is conceivable is to say that it is not ruled out *a priori* even upon ideal reflection. Moreover, also note that the relevant worlds here are those considered as *actual*. So, (Z1) tells us that it is *a priori* coherent, even after ideal reflection, that there are worlds, considered as actual, where $P \& \sim Q$ obtains.⁷

⁵ For the purposes of this paper, following Braddon-Mitchell, I will use the terms 'phenomenal consciousness' and 'qualia' interchangeably.

⁶ For instance, if phenomenally conscious experiences are required for the possession of phenomenal concepts, and that this possession bestows one with the *a priori* knowledge that one has had such experiences, we can't coherently employ these concepts to conceive of the actual world as a world that lacks consciousness. Nevertheless, we might still be able to coherently employ them to conceive of the actual world as a world that lacks some conscious states. I thank an anonymous referee for this point.

⁷ To consider a world as actual, roughly, is to suppose of a certain state of affairs (or epistemic scenario) that that state of affairs obtains in the actual world. (This is as opposed to supposing that that state of affairs obtains in some counterfactual world). So to claim that a world considered as actual is conceivable, in effect, is to say that that state of affairs isn't ruled out *a priori*. For more detail, see Chalmers (2006, 2010).

Premise (Z2) states that if such scenarios are conceivable, they are metaphysically possible. To say that a scenario is metaphysically possible is to say that there are worlds, considered as *counterfactual*, where the scenario obtains. So (Z2), in effect, tells us that if it is *a priori* coherent that there are worlds, considered as actual, where $P \& \sim Q$ obtains, then there are worlds, considered as counterfactual, where $P \& \sim Q$ obtains.

Premise (Z3) tells us that if there are worlds, considered as counterfactual, where $P \& \sim Q$ obtain, then physicalism is false. Physicalism is identical to, or entails, the thesis that the phenomenal truths at our world supervene on the physical truths at our world. This supervenience thesis is understood as a modal thesis, which entails that minimal physical duplicates of our world are also phenomenal duplicates of our world. $P \& \sim Q$ being true of some world considered as counterfactual contradicts this. Therefore, it is concluded that physicalism is false.

The conditional analysis of phenomenal concepts is supposed to undermine this argument. But what's more, it is supposed to be more advantageous than alternative ways of undermining the argument because it does so in way that lets us have the zombie intuition: the intuition that zombie worlds are conceivable. Its detractors, however, argue that the analysis fails in both these respects.

3. The Conditional Analysis

According to the conditional analysis of 'qualia', it is *a priori* true that:

CA1) If there are spooky states (SS) in the actual world, 'qualia' refers (rigidly) to SS.

CA2) If there are no SS in the actual world, 'qualia' refers (non-rigidly) to the occupiers of the relevant functional roles.

Here let 'spooky states' be a term for the nonphysical properties of the relevant type. We can think of these properties in turn as those properties, if there are any, which play the descriptive roles that dualists associate with our folk concept of qualia. So they are the properties which are responsible for the sorts of judgments we make when we say that we are in pain, feel nauseous, taste the rich tannins in fine wines,

and so on. Candidates for such properties include raw feels⁸, what-it-is-like for us to have experiences⁹, properties whose intrinsic essential natures are revealed to us in experience¹⁰ etc. Conditional (CA1) states that if the actual world has such properties, ‘qualia’ in all worlds, considered as actual and counterfactual, refers to them.

We can think of the functional roles as the roles physicalists associate with our concept of qualia. Presumably these roles will be causal roles. If you are an analytic functionalist, the functional roles will be decided by folk psychology, whereas if you are an empirical functionalist, the roles will be decided by the relevant empirical sciences. How to characterize these roles, as well as how to characterize spooky states, is an in-house debate we needn’t weigh in on.¹¹ What is relevant for our purposes is that (CA2) states that if the actual world lacks SS, ‘qualia’ in all worlds considered as actual refers to whatever occupies the relevant functional roles in these actual worlds. Moreover, ‘qualia’ in all worlds considered as counterfactual refers to the occupiers of the relevant functional roles in these counterfactual worlds.

To elaborate, functionalists suppose that functional roles are multiply realizable. Insofar as we suppose this, we shouldn’t take ‘qualia’ to refer rigidly to the actual occupiers of the functional roles. For example, even if it turned out that the actual world lacks SS, we shouldn’t take ‘qualia’ to refer rigidly to the actual physical properties that play these roles. This is because, provided the roles are multiply realizable, there would be worlds, considered as counterfactual, where these roles are occupied by properties alien to our world. Such worlds would also have qualia according to functionalists given that the relevant roles are occupied.

Two clarifications. First, even in the absence of SS in the actual world, we can still token identify ‘qualia’ with the physical properties that play the relevant functional roles in the actual world. This will have the consequence that while there could be worlds that have qualia in the absence of these physical properties, all worlds

⁸ See Tolman (1932).

⁹ See Nagel (1974).

¹⁰ See Johnston (1992).

¹¹ For an overview of these roles, see Braddon-Mitchell and Jackson (1997).

where these properties are instantiated and occupy the relevant functional roles will also be worlds that have qualia.¹²

Second, for the purposes of defending the conditional analysis, nothing hinges on whether we take ‘qualia’ to be non-rigid (on the assumption that the actual world lacks SS) except that we make the analysis more appealing to functionalists by doing so. If one is a physicalist who doesn’t allow for multiple realizability, one can take ‘qualia’ to refer rigidly to the actual occupants of the relevant physicalist roles. This too will have the consequence that all worlds in which the actual physical occupants are instantiated will also be worlds that have qualia.

The conditional analysis can also be formulated using the 2-dimensional semantic framework in which the conceivability of zombies is usually addressed. According to this framework, semantic expressions (or in this case concepts) are associated with two types of intensions: the primary and secondary intensions. The primary intension of a concept is a function from epistemic scenarios to extensions; on the supposition that the worlds considered are actual, whereas the secondary intension is a function from worlds to extensions; on the supposition that the worlds considered are counterfactual.

What is important to note is that the primary intension tracks what our concepts refer to in the actual world, whereas the secondary intensions track what our concepts refer to in counterfactual worlds based on the supposition that the reference of the concept in the actual world is already fixed. This supposition becomes significant when rigid concepts are involved. If a concept is not rigid, what it refers to in the actual world doesn’t determine what it refers to in counterfactual worlds.

¹² There could be worlds where the physical properties are instantiated but, due to blockers, don’t play the relevant functional roles. Such worlds won’t have qualia. E.g. see Hawthorne (2002) and Leuenberger (2008). However, the relevant worlds we are concerned with here are *minimal* physical duplicates of our world. Such worlds, by stipulation, won’t have blockers. So we can suppose that the instantiation of the relevant physical properties guarantees that they also play the relevant functional roles. Note, this might work as a response to Alter’s third objection as well.

Nevertheless, if the concept is rigid, what it refers to in the actual world will also determine what it refers to in any given counterfactual world.¹³

The proponents of the conditional analysis typically take it to be *a priori* true that the ‘qualia’ is rigid if the actual world has spooky states, but not so if the actual world lacks such states.¹⁴ Once we take this into account, we can reformulate the conditional analysis using the 2-dimensional framework. According to such a formulation, it is *a priori* true that:

CA1') If the actual world has SS, the primary and secondary intensions of ‘qualia’ track its extensions as SS.

CA2') If the actual world lacks SS, the primary and secondary intensions of ‘qualia’ track its extensions as the occupiers of the relevant functional roles.

The first conditional states that if the actual world has SS, the primary and secondary intensions of ‘qualia’ track its extension as SS. That is, in all worlds considered as actual and counterfactual, ‘qualia’ refers to SS. And the second conditional states that if the actual world lacks SS, the primary intension of ‘qualia’ tracks its extension as the actual occupants of the relevant roles, whereas the secondary intensions track its extension as the local occupants of the relevant functional roles. That is, in all worlds considered as actual, ‘qualia’ refers to the actual occupants of the relevant functional roles, whereas in any given counterfactual world, ‘qualia’ refers to whatever occupies the relevant roles in that world.

As is apparent, these conditionals carry no new information to that of the prior formulation. Therefore, a commitment to 2-dimensional semantics isn’t required to buy into the conditional analysis. All we need is a commitment to rigid concepts. In particular, a commitment to ‘qualia’ being rigid if the actual world were to have SS. Nonetheless, since the debate about the conceivability of zombie worlds is often conducted using the jargon of 2-dimensional semantics, the reformulated conditionals

¹³ For more on 2-dimensional semantics, see Chalmers (1996, 2006) and Jackson (1998). Note, Jackson calls the primary intension the ‘A-intension’, and the secondary intension the ‘C-intension’.

¹⁴ This issue will be taken up in §6.2.

are intended to help us see more clearly how the analysis enables us to undermine the conceivability argument in a way that allows for the zombie intuition.

4. The Zombie Intuition

The first premise of the conceivability argument tells us that zombie worlds are conceivable. This is controversial. Many, however, grant the zombie intuition: the intuition that zombie worlds are conceivable. Consequently, many find the conceivability argument persuasive, and take this to show that qualia are somehow over and above anything physical.

There are two distinct ways the conditional analysis can be used to undermine the conceivability argument whilst still accommodating the zombie intuition. The first way undermines the argument by allowing for the zombie intuition in virtue of zombie worlds themselves being conceivable. The second way undermines the argument by explaining the zombie intuition in virtue of zombie worlds being meta-conceivable.

4.1 The Conceivability of Zombie Worlds

The first way we can employ the conditional analysis to accommodate the zombie intuition can be formulated in terms of the following argument:

- 1a) If there are SS in the actual world, 'Q' in all worlds, considered as actual, refers to some SS.
- 2a) It is conceivable that there are SS in the actual world.
- 3a) Therefore, it is conceivable that there are worlds, considered as actual, where 'Q' refers to some SS.
- 4a) If there are worlds, considered as actual, where 'Q' refer to some SS, it is conceivable that there are worlds, considered as actual, where P&~Q obtains, as Q can fail to obtain in virtue of SS not being instantiated in these worlds.
- 5a) Therefore, it is conceivable that there are worlds, considered as actual, where P&~Q obtains.

Here (1a) is entailed by (CA1). If ‘qualia’ refers to spooky states on the supposition that the actual world has such states, then ‘Q’, which stands for some arbitrary quale, refers to some spooky state.

Premise (2a) tells us that the existence of some nonphysical substance, which is responsible for our phenomenal judgments, isn’t ruled out *a priori* even upon ideal reflection. The truth of this premise is supposed to be self-evident. Nonetheless, if this isn't altogether clear, consider the assumption behind it.

Spooky states, as we noted earlier, are the states that play the descriptive roles dualists associate with our concept of qualia. The proponents of the conditional analysis, I take it, concede that there are at least some such roles that are *a priori* coherent, and thereby that might be occupied in the actual world. Which in turn entails that it is conceivable that there are SS, of some form or another, in the actual world.

Note that this assumption isn’t without controversy, as there are physicalists, like Dennett (1991), who question the coherence of dualist notions of ‘qualia’. This isn’t the place to tackle the controversy. For now, let us concede that if all such notions were *a priori* incoherent, (2a) would turn out false. But let us also note two points in favour of the assumption by way of motivating (2a). First, even though some physicalists contest the coherence of certain dualist notions, a systematic account of how all such notions are incoherent is so far absent. And this is so regardless of whether charges of incoherence against specific dualist notions are justified, which is itself a controversial issue. Second, while some physicalists do deny the coherence of dualist notions, much of the significant disagreement between physicalists and dualists has concerned not whether it is *a priori* coherent that anything occupies the dualist roles, but rather whether anything *actually* occupies these roles.¹⁵ Reasons such as these, I take it, make (2a) plausible.

Premise (3a) follows straightforwardly from the conjunction of (1a) and (2a): if it is conceivable that there are SS in the actual world, and if it is true that provided

¹⁵ E.g. while Lewis (1995) denies that there are any states whose intrinsic essential natures are revealed to us in experience, this denial rests not on the *a priori* incoherence of this revelation thesis, but on the (supposed) falsity of the claim that these natures are *actually* revealed to us in experience.

there are SS, 'Q' refers to some SS, then it is conceivable that 'Q' in the actual world refers to some SS.

Premise (4a) is self-explanatory: if 'Q' in the actual world refers to some SS, it is conceivable that there are worlds, considered as actual, where P holds and Q doesn't because in such worlds Q could fail to obtain in virtue of lacking SS.

The conclusion (5a) follows from the conjunction of (3a) and (4a): if it is conceivable that 'Q' in the actual world refers to some SS, and if 'Q' referring in this way allows for the conceivability of worlds, considered as actual, where $P \& \sim Q$ obtains, then it is conceivable that there are worlds, considered as actual, where $P \& \sim Q$ obtains.

Now for the zombie intuition *vis-à-vis* the conceivability argument: the first premise of this argument tells us that zombie worlds are conceivable. That is, (Z1) states that it is conceivable that there are worlds, considered as actual, where $P \& \sim Q$ obtains. As we have seen, such worlds are conceivable: Q could fail to obtain in such a world in virtue of that world lacking SS. Therefore, the conditional analysis predicts that zombie worlds are conceivable; which in turn is to allow for the zombie intuition.

It is worth pointing out that there is a stronger way of construing zombie worlds, which is ruled out by the analysis. On this way, the first premise of the conceivability argument is as follows:

(Z1*) $PT \& \sim Q$ is conceivable.

Here 'T' is a *that's-all* clause or a totality fact, which tells us that the aforementioned truths, in this case, the physical ones, exhaust the truths about our world. (Z1*), then, states that we can conceive of a world, considered as actual, where P obtains, and that's all the truths that obtain in this world, and that this world still lacks Q.

The second conditional tells us that it is *a priori* true that if the actual world lacks SS, 'qualia', in all worlds considered as actual, refers to the occupiers of the relevant functional roles. More specifically, we saw that in such worlds, qualia would be type identical to the occupiers of the functional roles and token identical to the actual physical occupiers of these functional roles. Consequently, the analysis rules out the conceivability of scenarios, considered as actual, where PT obtains and Q

doesn't. If PT obtains, any given quale, in all worlds considered as actual, would be something physical; namely, some physical occupier of some relevant functional role.

4.2 The Meta-Conceivability of Zombie Worlds

The second way we can employ the analysis to accommodate the zombie intuition tries to explain the strong zombie intuition: the intuition that we can conceive of worlds, considered as actual, where $PT \& \sim Q$ obtains. To clarify, this approach doesn't allow for the conceivability of such worlds, but rather allows for their meta-conceivability, and argues that we often conflate the two. The strategy, formulated as an argument, runs as follows:

- 1b) If there are SS in the actual world, 'Q', in all worlds considered as counterfactual, refers to some SS.
- 2b) It is conceivable that there are SS in the actual world.
- 3b) Therefore, it is conceivable that there are worlds, considered as counterfactual, where 'Q' refers to some SS.
- 4b) If there are worlds, considered as counterfactual, where 'Q' refers to some SS, it is conceivable that there are worlds, considered as counterfactual, where $PT \& \sim Q$ obtains, as Q can fail to obtain in virtue of SS not being instantiated in these worlds.
- 5b) Therefore, it is conceivable that there are worlds, considered as counterfactual, where $PT \& \sim Q$ obtains.

This argument works in an analogous way to the one we just explored. The difference is that here we are concerned with counterfactual worlds, and therefore the realm of metaphysical possibility, as opposed to worlds considered as actual and the realm of epistemic possibility.

As before, (1b) is entailed by (CA1), whereas (2b) is the same as that of the previous argument. And (3b) follows from the conjunction of these two premises: if it is conceivable that there are SS in the actual world, and if it is true that provided there are SS in the actual world, 'Q' refers rigidly to some SS, then it is conceivable that there are worlds, considered as counterfactual, where 'Q' refers to some SS.

Premise (4b), like (4a), is self-explanatory: if ‘Q’ in all counterfactual worlds refer to some SS, it is conceivable that there are counterfactual worlds where $PT \& \sim Q$ obtains, as Q can fail to obtain in virtue of SS not being instantiated in these worlds.

Finally, (5b) follows straightforwardly from the conjunction of (3a) and (4b): if it is conceivable that ‘Q’, in all worlds considered as counterfactual, refers to some SS, and if ‘Q’ referring in this way allows for the conceivability of worlds, considered as counterfactual, where $PT \& \sim Q$ obtains, then it is conceivable that there are worlds, considered as counterfactual, where $PT \& \sim Q$ obtains.

Now recall that if there are worlds, considered as counterfactual, where a given scenario obtains, we take this to be constitutive of that scenario being metaphysically possible. So (5b), in effect, tells us that it is conceivable that $PT \& \sim Q$ worlds are metaphysically possible. This helps us explain the strong zombie intuition: as we noted earlier, insofar as (CA2) is true *a priori*, it would not be *a priori* coherent that there are worlds, considered as actual, where $PT \& \sim Q$ obtains. Nonetheless, according to the strategy just spelt out here, while $PT \& \sim Q$ worlds aren’t conceivable, they are meta-conceivable, i.e. it is conceivable that they are metaphysically possible. Therefore, given that intuitions don’t tend to be very fine-grained, this allows us to explain away the strong zombie intuition by claiming that we confuse the *a priori* coherent intuition that strong zombie worlds are meta-conceivable with the *a priori* incoherent intuition that strong zombie worlds are conceivable.

5. Undermining the Conceivability Argument

There are two ways to undermine the conceivability argument; each corresponding to the two different ways we explain the zombie intuition.

5.1 Deny Premise (Z2)

According to the first way we undermine the argument, we grant that zombie worlds, understood as worlds considered as actual where $P \& \sim Q$ obtains, are conceivable, but we deny that this entails their metaphysical possibility. We can formulate this strategy as follows:

- 1c) (CA1) and (CA2) are true.

- 2c) If (CA1) and (CA2) are true, zombie worlds are only metaphysically possible if there are SS in the actual world.
- 3c) Zombie worlds are conceivable. [Regardless of whether or not there are SS in the actual world].
- 4c) Therefore, in the case where we conceive of zombie worlds when the actual world lacks SS, we would be conceiving of a scenario that is metaphysically impossible.
- 5c) Therefore, the conceivability of zombie worlds doesn't entail their metaphysical possibility.

Premise (1c) is entailed by the conditional analysis, and should be clear. Premise (2c), however, is in need of explanation.

The conditionals of the analysis have implications for which scenarios are and aren't metaphysically possible: (CA1) tells us that if there are SS in the actual world, 'Q' refers rigidly to some SS. Therefore, the analysis tells us that if there are SS in the actual world, there would be worlds, considered as counterfactual, where $P \& \sim Q$ obtains because Q could fail to obtain in virtue of such worlds lacking SS. In other words, if there are SS in the actual world, zombie worlds are metaphysically possible.

In contrast, (CA2) tells us that if the actual world lacks SS, 'Q' refers to some occupier of some relevant functional role. As we noted earlier, we can, therefore, token identify Q with some actual physical occupant of a relevant role in the actual world. Therefore, the analysis tells us that if the actual world lacks SS, there would be no worlds, considered as counterfactual, where $P \& \sim Q$ obtains because if P obtains, then so does Q. In other words, if there are no SS in the actual world, zombie worlds are metaphysically impossible. Overall, then, the analysis tells us that zombie worlds are only metaphysically possible if the actual world has SS.

We already saw an explanation for premise (3c) in §4.1. Premise (2c) in conjunction with (3c) tells us that we can conceive of scenarios that are metaphysically impossible. In particular, it tells us that while $P \& \sim Q$ is conceivable of a world considered as actual, in the case where the actual world lacks SS, we would in fact be conceiving of a scenario that is metaphysically impossible. And this in turn

allows us to falsify (Z2), as we see that the conceivability of zombie worlds doesn't entail their metaphysical possibility.

A proviso. The falsity of (Z2) needn't be construed as a rejection of the useful rule of inference that conceivability is a good guide to metaphysical possibility. Insofar as conditional concepts are rare, which they appear to be, this inference could still hold. All we need to do is be cautious that this inference doesn't hold up to scrutiny when such concepts are in play. In this case, we see that since the concept 'qualia' is conditional, just because we can conceive of minimal physical duplicates of our world that lack qualia, this doesn't ensure that such worlds are metaphysically possible.

5.2 Deny Premise (Z1*)

According to the second way we undermine the conceivability argument, we stipulate that zombie worlds are worlds, considered as actual, where $PT \& \sim Q$ obtains. Once we make this stipulation, we see that the conceivability argument runs into trouble straight away. Insofar as we grant that (CA2) is true *a priori*, we see that we can't conceive of zombie worlds so construed. That is, if it is true *a priori* that if the actual world lacks SS, 'Q' in the actual world refers to some physical occupier of a relevant functional role, then we can't conceive of worlds, considered as actual, where physical properties occupy the relevant functional roles, and there are no SS, but there is no Q. Therefore, the conceivability argument is undermined because (Z1*) turns out to be false.

6. Objections

There are two objections to the conditional analysis that are supposed to be devastating. The first contests the claim that the analysis explains the zombie intuition, and the second contests the claim that the analysis undermines the conceivability argument.

6.1 *The Wrong Intuition*

The first objection, which is raised by Chalmers (2005) and Alter (2007), states that the analysis doesn't explain the correct zombie intuition and, moreover, that it can't because the analysis is inconsistent with this intuition.

To elaborate, we saw that there are two ways to account for the zombie intuition. First, we might demonstrate how the conditional analysis entails that zombie worlds – worlds considered as actual where $P \& \sim Q$ obtains – are conceivable. As far as I can tell, both Hawthorne and Stalnaker take this approach. Second, we might show that strong zombie worlds – worlds considered as actual where $PT \& \sim Q$ obtains – while not being conceivable are meta-conceivable. And we might then argue that we conflate the conceivability of strong zombie worlds with their meta-conceivability. Braddon-Mitchell favours this approach.

In response, Chalmers and Alter object that the correct zombie intuition is the intuition that strong zombie worlds are conceivable. And on this basis, they argue that none of the proponents of the analysis account for this intuition. Moreover, that none of them can account for the intuition either because, as we saw earlier in §4.1 and §5.2, this intuition contradicts (CA2) being *a priori* true.

Before we quibble over whether the conditional analysis can be used to account for the correct zombie intuition, let us be perfectly clear that there are three potential candidates for this intuition:

- i) $P \& \sim Q$ is conceivable of worlds considered as actual.
- ii) $PT \& \sim Q$ is conceivable of worlds considered counterfactual.
- iii) $PT \& \sim Q$ is conceivable of worlds considered as actual.¹⁶

The first intuition, (i), is what I have regarded as the zombie intuition, whereas (ii) is the intuition that strong zombie worlds are meta-conceivable, and (iii) is the intuition that strong zombie worlds themselves are conceivable. I have labeled (iii) the strong zombie intuition but note that I use these labels merely as shorthand by which to refer

¹⁶ The obvious fourth alternative – ' $P \& \sim Q$ is conceivable of worlds considered as counterfactual' – is unmotivated, as no one supposes that this is the relevant zombie intuition. Moreover, even if this alternative were the relevant intuition, like (ii), it wouldn't undermine physicalism given that physicalism is a thesis about the actual world.

to the relevant intuitions, and by no means want to suggest that the correct zombie intuition is (i). What I do want to stress, however, is that all three of these intuitions are *prima facie* plausible candidates for the zombie intuition we have.

A close reading of Chalmers (1996) suggests that the intuition he has in mind, and the one a lot of us subsequently signed up for arguably, is (i). It is this intuition, for instance, that both Stalnaker and Hawthorne try to show is compatible with physicalism. Moreover, it might also turn out that the intuition we have is (ii). As we noted earlier, intuitions aren't very fine-grained. Consequently, we can follow Braddon-Mitchell and argue that we confuse the conceivability of $PT \& \sim Q$ with its meta-conceivability; i.e. with (ii). Finally, Chalmers (2005) and Alter (2007) are explicit that the intuition they have is (iii). Therefore, on testimony, we should grant that (iii) might also, *prima facie*, be the zombie intuition; at least as had by some of us.

Once we see that there are three distinct and *prima facie* plausible candidates for the zombie intuition, how to respond to the first objection depends on which intuition is at issue. As we saw earlier, the conditional analysis allows us to grant intuitions (i) and (ii). Therefore, insofar as it is either (i) or (ii) which is the intuition that most of us have, the conditional analysis can be said to not only enable us to undermine the conceivability argument but to do so in a way that is compatible with our having the zombie intuition.

If, however, the zombie intuition at issue is (iii), the conditional analyst can respond by claiming that the analysis enables us to undermine the conceivability argument in a way that explains away the zombie intuition: since (CA2) being true *a priori* contradicts $PT \& \sim Q$ being *a priori* coherent of worlds considered as actual, the proponent of the analysis, as we saw, can reject the conceivability argument based on (Z1*) being false. Moreover, the proponent can then point to either intuition (i) or (ii) to explain why we mistakenly take ourselves to have the intuition that $PT \& \sim Q$ is conceivable of worlds considered as actual.

This way of responding is unlikely to please the objectors, as in this case the proponent of the analysis ends up denying an intuition that the objectors claim to have. Nonetheless, in defense consider the three following points. First, neither objector says anything to deny that the zombie intuition we have is (i) besides

asserting that the correct zombie intuition is (iii). Therefore, insofar as we take account of how our intuitions aren't very fine-grained, and allow for fallibility with regards to our introspective judgments concerning these intuitions, it still remains plausible that the intuition the objectors actually have is (i).

Second, we have no good grounds to deny the plausibility of our conflating (ii) with (iii) either. In all fairness, Alter does attempt to discredit the plausibility of this conflation. In brief, while he concedes that we sometimes conflate incoherent modal claims with coherent doubly modal claims, he argues that, in such cases, once the distinction between the two are pointed out to us, we lose the inclination to find the incoherent modal claims themselves conceivable. For example, he argues that once we make Kripke's (1972) distinction between *imagining Hesperus without itself* (an incoherent scenario) and *Hesperus not being the morning's last visible heavenly body* (a coherent scenario), we lose the inclination that we can imagine or conceive of the former scenario. But, argues Alter, this is not so with (ii) and (iii): even when the distinction between them is pointed out, "it does not weaken our confidence in the coherence of the [strong] zombie-world hypothesis" (2007: 245-246). And from this, he concludes that it is unlikely that the intuitive force of the strong zombie intuition derives from conflating (ii) with (iii).

In response, first note that the analogy between conflating (ii) with (iii) and conflating modal claims with doubly modal claims more generally is not only an analogy a proponent of the conditional analysis can accept, but one that he ought to encourage if he is to explain why we mistakenly take ourselves to have the strong zombie intuition. Nevertheless, the proponent ought to question Alter's justification for the proposed asymmetry. Once again, the justification amounts to nothing more than an assertion that, even once the distinction is made, we find strong zombie worlds conceivable. Therefore, insofar as we allow for the aforementioned fallibility concerning our intuitions, this assertion alone won't suffice to undermine the hypothesis that we think we can conceive of (iii) because we conflate it with the *a priori* coherent scenario (ii).

Third, and most crucially, dialectically, there is no onus on the proponent of the conditional analysis to allow for the strong zombie intuition, as it concerns conceiving of a scenario which the proponent takes to be conceptually incoherent. To

elaborate, the proponents themselves take (CA2) to be true *a priori*, which means that they must find strong zombie worlds *a priori* incoherent. Therefore, they wouldn't be moved by any demands made by the objectors to accommodate the strong zombie intuition. The only way for the objectors to proceed, then, is to show that (CA2) isn't true *a priori*, which brings us to the second objection.¹⁷

6.2 Apriority

The second main objection to the analysis, also raised by Chalmers and Alter, states that (CA2) isn't true *a priori*. This objection is potentially devastating because, were this claim true, not only would the conditional analysis fail to provide a conceptual analysis of 'qualia', it would also subvert the two ways the analysis can be employed to undermine the conceivability argument.

Very briefly, if (CA2) isn't *a priori*, we can no longer hold that strong zombie worlds, i.e. world considered as actual where $PT \& \sim Q$ obtains, is *a priori* incoherent. Consequently, we can no longer deny (Z1*). But also note that, if (CA2) isn't *a priori*, we can't know *a priori* that zombie worlds, i.e. worlds considered as actual where $P \& \sim Q$ obtains, are metaphysically impossible if it turns out that the actual world lacks SS. Thus, we no longer have any *a priori* grounds for claiming that the conceivability of zombie worlds doesn't entail their metaphysical possibility. Consequently, we can no longer undermine the conceivability argument by contesting (Z2) either.

So objection two proves to be potentially devastating, but why believe it? First, consider the reasons we have for accepting that (CA2) is true *a priori*. According to the proponents of the analysis, we have grounds to believe that (CA2) is true *a priori* because it explains how agents would and do behave. For instance, Hawthorne argues that even if an oracle told you that the actual world lacks SS, you would "remain convinced that you do feel pain sometimes and will reckon as pain

¹⁷ Haukioja (2008) also provides us with a response to the first objection, which involves exploiting the absence of an appearance/reality distinction when it comes to qualia to show that strong zombie worlds are *a priori* incoherent (see pg. 150). This response, however, won't work as Haukioja ends up relying on introspective cum *a posteriori* evidence about our own phenomenal states to rule out such worlds.

whatever plays the pain role” (2002: 26). Similarly, Braddon-Mitchell points out that the philosophers amongst us who once believed in dualist conceptions of qualia and then became physicalists aren’t typically error theorists about qualia but, instead, attempt to give physicalist accounts of it.

Here it is worth spelling out the relationship between our behavior and apriority, as this isn’t altogether clear. For Braddon-Mitchell, there are two methods by which we can have access to the meanings of our concepts: introspection; by consulting our own understanding of our concepts, and empirical inquiry; by seeing what we take to be the referents of our concepts based on certain relevant discoveries about the actual world. Behavioural evidence links up to apriority because it falls under the latter category. Since we wouldn’t be error theorists about qualia even were we to discover that the actual world lacks SS, this is supposed to, then, be evidence that it is an aspect of the meaning of our concept that were the actual world to lack SS, ‘qualia’ would refer to the occupiers of the relevant functional roles. In other words, this is supposed to be evidence that (CA2) is true *a priori*.¹⁸

There are two ways to deny that (CA2) is true *a priori* in light of this evidence. First, one might appeal to intuitions to the contrary. Chalmers (2005), for example, argues that the conditional isn’t *a priori* because it contradicts both the (strong) zombie intuition and Jackson’s (1982) Mary intuition. This response, however, proves inadequate. These intuitions do nothing to directly undermine the behavioural evidence itself. Instead, all they do is show that we tend to have conflicting beliefs about worlds, considered as actual, in which PT&~Q obtains when these beliefs are individuated introspectively and behaviourally.

But more worryingly, appealing to these intuitions in this context also arguably begs the question. Supposing the (strong) zombie intuition, arguably, begs the question against (CA2) being true *a priori*. To elaborate, what’s at issue here is the *a priori* status of (CA2). If (CA2) were *a priori* true, the strong zombie intuition,

¹⁸ This account of behavioural evidence as a guide to apriority isn’t without controversy. This is presumably why Braddon-Mitchell himself qualifies that such evidence makes the conditional “in some sense *a priori* and analytic” (2003: 130) That said, since neither Chalmers nor Alter take issue with this directly, and since addressing it would take us far too further afield, we can leave this controversy aside.

as we noted in §5.2, would be *a priori* incoherent. The converse, of course, also holds true. That is, if the strong zombie intuition were *a priori* coherent, (CA2) wouldn't be *a priori* true. However, simply stating the converse without argumentation, which is what Chalmers appears to do here, would beg the question against the *a priori* status of (CA2). (An opponent of the conditional analysis might object that assuming (CA2) is *a priori* true is also to beg the question against the strong zombie intuition. However, this isn't so because the proponents of the analysis appeal to behavioural evidence to support their claim).

The same point can be made with regards to the Mary intuition. The conditional analysis can be used to undermine Jackson's knowledge argument in ways analogous to the two ways we can use it to undermine the conceivability argument. It is beyond the scope of this paper to defend this claim. However, for present purposes, we can note that the intuition which would undermine the *a priori* status of (CA2) would be the intuition that the actual world is such that (i) Mary knows all the physical truths at this world, (ii) these truths exhaust all the truths at this world, and yet (iii) she knows something new when she has a novel experience of a red object. Now, nowhere in Jackson (1982) does he mention claim (ii), so the intuition we need here is a distinct Mary intuition to the one a lot of us claim to share. Consequently, the opponent of the analysis needs to say more to motivate this new intuition. And failing this, to simply appeal to it would be to beg the question.¹⁹

The second way to deny the *a priori* status of (CA2) in light of behavioural evidence is to try to directly undermine this evidence itself. The proponents of the analysis argue that (CA2) is true *a priori* because it explains how agents behave, but one can argue, as Alter does, that agents behave in the way they do because of *a posteriori* knowledge of their conscious experiences. For example, suppose a group of agents were told by an oracle that there are no SS in the actual world. These agents would still behave as if there were pains, but this is really "because of their

¹⁹ The conditional analysis allows for the original Mary intuition in a way analogous to how it allows for the zombie intuition in §4.1. Similarly, it allows us to block any inferences made from the Mary intuition to the falsity of physicalism in a way analogous to how we use it to block the inference from conceivability to metaphysical possibility in §5.1. In particular, we block the inference that if $P \rightarrow Q$ is not *a priori*, $P \rightarrow Q$ is not necessary. See Chalmers (2005).

confidence in their *a posteriori* knowledge that pains occur” (Alter 2007: 240). This response proves more effective, as it provides us with an explanation that is both *prima facie* plausible and which directly seeks to undermine the evidence we have for supposing that (CA2) is true *a priori*.²⁰

Here it is worth stressing a point that both the proponents and opponents of the analysis fail to mention: whether (CA2) is true *a priori* with regards to our actual folk concept is something which we can know only empirically. To clarify, this isn’t to deny that we have *a priori* access to the meanings of our own concepts. The point, rather, is while I may *a priori* know the meanings of my own concepts, I can’t know *a priori* the meanings of the folk concepts at large. The only way I can find this out is to poll the folk. In this vein, while I may know *a priori* that my concept ‘qualia’ has it that if the actual world lacks SS, ‘qualia’ refers to occupiers of the relevant functional roles, I can’t know *a priori* that this holds true of the folk concept. And neither can any of the proponents or opponents of the conditional analysis.

Accordingly, we will have to refrain from settling the current debate until we do some empirical polling. That said, in the face of pressing disagreement, let me explain why I think the interlocutors themselves aren’t in a position to claim that (CA2) is true *a posteriori*. The explanation will be easier if we focus on the 2-dimensional reading of (CA2), which was formulated as follows:

(CA2') If the actual world lacks SS, the primary and secondary intensions of ‘qualia’ track its extensions as the occupiers of the relevant functional roles.

To recap, this conditional tells us that if the actual world lacks SS, then the primary intension of ‘qualia’ tracks its extension as the actual occupiers of the relevant functional roles, and the secondary intensions track its extensions as the local occupiers of the relevant roles. What is at issue *vis-à-vis* the second objection is what

²⁰ Haukioja argues that while (CA2) isn’t *a priori*, once we modify it to make it explicit that we are talking about *our* sensations and *our* physical or nonphysical states, the conditional will be true *a priori* (see pg. 148). Both Chalmers and Braddon-Mitchell suppose that the actual worlds mentioned in the conditionals are really centered worlds, which entails the modification Haukioja suggests. Therefore, the modification itself is unlikely to convince the interlocutor. What’s more, it still leaves room for the interlocutor to argue that we know that the modified conditional is true *a posteriori*.

the primary intension of ‘qualia’ tracks as its extension. The advocates of the analysis argue that it tracks the actual occupiers of the functional roles, whereas its detractors claim that it won’t track anything. That is, if the actual world lacks SS, ‘qualia’ would fail to refer.

Now, we can note, following Chalmers (1996: 57), that the primary intension of a given concept is independent of empirical factors because it *specifies* how the reference depends on the way the actual world turns out, as opposed to depending on the way the actual world itself turns out. Insofar as this is true, for any given concept, what its primary intension tracks as its extension, on the supposition that the actual world turns out to be in a such and such way, must be known *a priori*. Thus, on the supposition that there are no SS in the actual world, whether or not the primary intension of ‘qualia’ tracks its extension as whatever plays the relevant functional roles, must also be known *a priori*. Accordingly, if (CA2') is true, it must be true *a priori*.

The objectors could contest this. But, in doing so, they would face the following dilemma. On the one hand, they could insist that we know what the primary intension of our concept ‘qualia’ tracks as its extension, on the supposition the actual world is thus and so, *a posteriori*. But this would be to commit a conceptual error. Empirical information, as we have just seen, can’t technically play a role in telling us what the primary intension of a concept tracks as its extension.

A point also worth noting is that while empirical information plays no role in determining what the primary intension of ‘qualia’ picks out as its reference, such information can, and generally does, play a role in determining what the secondary intensions pick out. To elaborate, secondary intensions fix the reference given whatever world is in fact the actual world. So empirical information about the way the actual world turns out can play a role in determining what the secondary intensions of a concept help pick out.²¹ Maybe part of the confusion, then, is due to the failure of the interlocutors to realize that what they are contesting when they contest the apriority of the second conditional is really a claim about what the primary intension picks out. This is one reason why it helps to cash out the conditionals in such a way

²¹ See Jackson (1998) for more on why secondary intensions tend to be *a posteriori* whereas the primary intensions are *a priori*.

that we make the claims about primary intensions explicit. Once we cash out the second conditional in this way, we see that it can't be true *a posteriori*.

On the other hand, since we don't know what the primary intension of 'qualia' tracks as its extension *a posteriori*, if the objectors deny that we can know it *a priori* either, they have to acknowledge that we can't know what the primary intension of 'qualia' tracks as its extension. This simply won't do. Everyone in the present discourse agrees that if there are no SS, qualia are physical states. So everyone appears to grant, be it explicitly or tacitly, that on the supposition that the actual world lacks SS, the primary intension of 'qualia' tracks its extension as physical states. The debate, let me stress, is over whether the second conditional is *a priori*, not whether its true. Therefore, it turns out that neither horn is acceptable. Consequently, we can conclude that insofar as we grant that (CA2') is true, we need to grant that it is true *a priori*.

So here we see that while we can't completely settle the issue of whether the second objection works until we have the relevant empirical data, at present, we have good grounds to think that it doesn't work. Moreover, since we saw earlier that the first objection works only if the second does, we thereby have good grounds to think that the first objection doesn't work either. The overall lesson, then, is that neither objection proves to be devastating.

7. Conclusion

There are two distinct ways the conditional analysis can be used to undermine the conceivability argument in a way that accommodates the zombie intuition. If the zombie intuition is the intuition that it is conceivable that there are worlds, considered as actual, where $P \& \sim Q$ obtains, the analysis allows for this intuition because it predicts that such worlds are actually conceivable. The analysis, however, also shows that the conceivability of such worlds doesn't entail their metaphysical possibility, which allows us to reject the second premise of the conceivability argument that assumes this entailment.

In contrast, if the zombie intuition is the intuition that there are worlds, considered as actual, where $PT \& \sim Q$ obtains, the analysis doesn't allow for this

intuition, as it predicts that such worlds are inconceivable. Here we can reject the first premise of conceivability argument, which states that such worlds are conceivable. Moreover, since the analysis predicts that such worlds are meta-conceivable despite their inconceivability, we can accommodate the zombie intuition by arguing that we conflate the meta-conceivability of such worlds with their conceivability.

Once we note these two distinct ways the analysis can be used to undermine the conceivability argument, we are able to see that neither of the two main objections to the analysis proves devastating. The upshot, then, is that the conditional analysis remains a viable and attractive way of undermining the conceivability argument against physicalism.²²

The University of Otago

Raamy@yahoo.com

References

- Alter, T (2007). ‘On the Conditional Analysis of Phenomenal Concepts’. *Philosophical Studies*, 134: 235-253.
- Braddon-Mitchell, D and Jackson, F (1997/2006). *The Philosophy of Mind and Cognition*. Blackwell Publishing.
- Braddon-Mitchell, D (2003). ‘Qualia and Analytic Conditionals’, *The Journal of Philosophy*, 100: 111-135.
- Chalmers, D (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- (2005). ‘The Two-Dimensional Argument Against Materialism’, abridged version in B. McLaughlin (ed.) *Oxford Handbook of the Philosophy of Mind*. Oxford University Press, 2009.

²² I am grateful to Torin Alter, Ben Phillips, and two anonymous referees for this journal for their helpful comments on the draft.

- (2006). 'Two-Dimensional Semantics', in E. Lepore & B. Smith (eds.), *Oxford Handbook of the Philosophy of Language*. Oxford University Press.
- (2010). 'The Nature of Epistemic Space', in A. Egan and B. Weatherson (eds.), *Epistemic Modality*. Oxford University Press.
- Dennett, D (1991). *Consciousness Explained*. Boston: Little, Brown, and London.
- Haukioja, J (2008). 'A defence of the conditional analysis of phenomenal concepts', *Philosophical Studies*, 139: 145-151.
- Hawthorne, J (2002). 'Advice for Physicalists', *Philosophical Studies*, 108: 17-52.
- Jackson, F (1982). 'Epiphenomenal Qualia', *Philosophical Quarterly*, 32: 127-36.
- (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Clarendon Press.
- Johnston, M (1992). 'How To Speak of the Colours', *Philosophical Studies*, 68: 221-263.
- Kripke, S (1972). 'Naming and Necessity', in G. Harman and D. Davidson (eds.), *The Semantics of Natural Language*, Dordrecht: Reidel. Reprinted as *Naming and Necessity*, Cambridge, MA: Harvard University Press, 1980.
- Leuenberger, S (2008). 'Ceteris Absentibus Physicalism', *Oxford Studies in Metaphysics*, 4: 145-70.
- Lewis, D. K (1995). 'Should a Materialist Believe in Qualia?', *Australasian Journal of Philosophy*, 73: 140-144.
- Nagel, T (1974). 'What it is like to be a bat?', *Philosophical Review*, 4: 435-450.
- Stalnaker, R (2002). 'What is it like to be a zombie?', in T. S. Gendler & J. Hawthorne (eds.), *Conceivability and Possibility*. Oxford: Oxford University Press.
- Tolman, E.C. (1932). *Purposive Behavior in Animals and in Man*. New York: The Century Company.