Undergraduate Philosophy Journal of Australasia www.upja.online
Vol. 4, December 2022

Two-Dimensional Modal Semantics and the Zombie Intuition

Charlotte Senior[1]

The University of Sydney

Abstract

In this paper I discuss David Chalmers' 'zombie argument' in favour of a dualist theory of qualia. I begin by explaining Chalmers' original argument and David Braddon-Mitchell's use of Two-Dimensional Modal Semantics in "Qualia and Analytic Conditionals" (2003) to distinguish two types of conceivability and put pressure on the zombie intuition as proof of dualism. I then critically evaluate Chalmers' defence of the zombie intuition, focussing firstly on his reliance on microphysical properties and secondly on the questionable alignment between his deductive argument for the 'secondary conceivability' of philosophical zombies and actual experience. Stephen Yablo's account of conceivability, specifically regarding undecidability, is fruitful when explicating the latter issue. I end by concluding that Braddon-Mitchell's two-dimensional account can make sense of the primary conceivability of philosophical zombies, and Chalmer's defence of the zombie intuition at least partially relies on problematic argumentation. However, there are further problems elicited by Chalmers' paper surrounding which

[1] Charlotte graduated last year from the University of Sydney with an Arts Degree (double major in Philosophy and Anthropology), with honours in Philosophy. Her primary areas of interest are epistemology and metaphysics, of which she took a particular interest in during her final year of study. Charlotte's honours thesis focused on hyperintensionality and its relation to theories of impossible worlds.

structures necessitate qualia, meaning that the problems relating to physicalism his defence presents still bear significant weight.

1. Introduction

The zombie intuition is the purported conceivability of an exact physical copy of yourself that lacks conscious experience. This physical copy is utterly identical to you in terms of physical composition, meaning that the placement of every atom (or your preferred physical fundamental) is identical, and nothing has been added beyond your exact composition of atoms. The zombie intuition arises from the "epistemic gap" dilemma, whereby there is no full explanation of how conscious experience, or qualia, arises from physical brain states.[2] To be clear, the "epistemic gap" dilemma does not posit that we cannot theoretically give a full explanation of all the physical facts that *produce* qualia, but rather that we cannot comfortably say these physical facts *constitute* qualia. Frank Jackson's black and white room thought experiment illuminates the point well: if qualia is nothing over and above a collection of physical facts, then the physical facts associated with a certain qualitative experience (for example, seeing red), should constitute the experience itself.[3] At the very least, this is the most direct conclusion to draw if physical facts are one and the same with qualia, which in turn seems to be straightforwardly derivable from the proposition that physical matter is all that exists.

It is this proposition which is at the core of *physicalism*, which I shall use interchangeably with *materialism*. Both state the thesis that physical facts, or matter, is all that exists and thereby involved with the production and constitution of qualia. Physicalism, or materialism, is in opposition with dualism, which ultimately posits there is *something more* than physical facts in the production and constitution of qualia.

---

[2] Chalmers, David (2009) 'The Two-Dimensional Argument Against Materialism', in A. Beckermann, B.P. McLaughlin, and S. Walter, eds. *The Oxford Handbook of Philosophy of Mind*: 313. Oxford University Press.

[3] Jackson, Frank (1986) 'What Mary Didn't Know', *The Journal of Philosophy* **83**: 291-95.

I intend to focus on the analytic functionalist account of qualia as the core physicalist opponent to dualism. Thus, before I proceed, it is necessary to outline analytic functionalism as a thesis and clarify its place in the following arguments. The view holds that qualia is identical to the physical thing(s) or process(es) that plays the relevant *functional role*; this is what qualia is by definition.[4] The view takes this identity relation at face value and in that sense is the most straightforward version of the physicalist thesis to contrast to dualism. Analytic functionalism places no weight on exactly *what* physical facts constitute qualia, just that qualia is by definition whatever physical phenomena plays the right functional role(s).[5]

Because it is merely the functional role that matters, different qualitative experiences (for example, the experience of seeing red and that of tasting chocolate) could have completely different physical facts constituting them. Hence, it is not necessary for there to be anything consistent underlying all conscious experience.[6] Further, philosophers such as David Lewis (1990) and Lawrence Nemirow (1990) have maintained that qualia is whatever physical phenomena plays the right functional role whilst adding that the *visual representation* of red and all of the *physical facts* that define the subjective experience of seeing red are different 'modes of presentation' of red, whilst both being identical with the subjective experience of red.[7] I mention these details to elucidate the focus. These concerns are supplementary to the foundational claim that qualia is identical with the physical facts that play the relevant functional role. It is in this capacity that analytic functionalism will henceforth represent physicalism and oppose dualism.

---

[4] Braddon-Mitchell, David (2003) 'Qualia and Analytical Conditionals', *The Journal of Philosophy* **100**: 120.

[5] So, there could be vastly physical facts accounting for different qualitative experiences.

[6] Of course, this may still be the case contingently.

[7] See: Lewis, David (1990) 'What Experience Teaches', in W. Lycan ed. *Mind and Cognition: A Reader*: 499-515.

Dualists have used the "epistemic gap" as the foundation from which to introduce the conceivability argument against materialism. The argument runs thus: firstly, you can conceive of a philosophical zombie in the sense described earlier. Secondly, if something is conceivable, then it is metaphysically possible. (Henceforth, metaphysical possibility will be defined according to possible worlds. That is, a proposition is metaphysically possible if it is true in at least one possible world, where a possible world is a complete explanation of a way the world could have been. Everything that is metaphysically possible is logically possible, however there are propositions that are logically possible but not metaphysically possible.[8] Regrettably, brevity does not permit a further discussion on the consistency of this term within this paper.) Thirdly, if a physical duplicate of myself that lacks conscious experience is metaphysically possible, then I must have something non-physical accounting for my conscious experience. Therefore, materialism is false.

Chalmers (2009) has formalised this argument. Let Q be the fact that someone has qualia. Let P be the conjunction of all material facts about the world. For instance, P could include the proposition "the door to my house is red". This proposition can also be analysed in terms of more granular physical facts, perhaps about the chemical proposition of the paint, or even the placement of each and every atom (or physical fundamental). What matters is these facts detail the physical state of the world, and making up P, account for all physical phenomena.

The argument runs thus:

(1) P&~Q is conceivable

(2) If P&~Q is conceivable, P&~Q is metaphysically possible.

(3) If P&~Q is metaphysically possible, materialism is false.

[8] "The car is red but not coloured" is an example of a proposition that is logically possible but not metaphysically possible.

(4) Therefore, materialism is false.[9]

The most controversial of these premises is the second. Philosophers have challenged the idea that conceivability entails metaphysical possibility, particularly Yablo (1993) who gives a fruitful definition of conceivability that escapes the binary between conceivability and inconceivability, and Braddon-Mitchell (2004), who targets the zombie intuition and explains its existence using two-dimensional modal semantics. Chalmers (2009) defends the second and third premise by claiming it is valid to infer possibility from conceivability in the case of the zombie intuition, because this intuition is a special class of conceivability.

In this paper, I will start by outlining Braddon-Mitchell's account for the zombie intuition in terms of two-dimensional semantics. I will then sketch out Chalmers' response that rests on the notion that Braddon-Mitchell has explained the wrong kind of conceivability, and the right class of conceivability regarding the zombie intuition can be deductively proven. I will then critically analyse the course Chalmers takes to reach his conclusion, specifically his distinction between and use of 'intrinsic' and 'structural' profiles. Finally, I will critique the implication from this argument that we have so called 'secondary conceivability' of philosophical zombies and argue that this conclusion does not align with direct experience.

2. Two-Dimensional Account of the Zombie Intuition

In his paper "Qualia and Analytic Conditionals", David Braddon-Mitchell gives an explanation for the zombie intuition that does not undermine materialism. The basis of the argument is that conceivability and possibility claims are always made on the basis of what we believe is actual. Thus, our assessment of a world considered as counterfactual, in terms of the properties it possesses, is largely dependent on what we believe the actual world is like. If the actual world is one in which qualia is

[9] Chalmers, 'The Two-Dimensional Argument Against Materialism', 314.

non-physical and dualism is true, then it is an *a posteriori* necessity that qualia is non-physical, meaning this is both necessary and empirically discoverable[10]

Perhaps the notion of *a posteriori* necessity is best explained with an example using a physical phenomenon we encounter everyday: water. We know that water is H^2O and would classify this as a necessary truth, because there is no other chemical composition that could be water. Yet, we did not arrive at the truth that water is H^2O via reasoning (as we do with mathematical truths, for instance), but through experimentation and more broadly, *experience*. The role of experience, rather than pure reason, in uncovering this necessary truth makes it an *a posteriori* necessity that water is H^2O. Similarly, if scientists discovered a full explanation of qualia as physical brain states, the proposition "qualia are physical brain states" would be an *a posteriori* necessity.

The notion that we can empirically uncover a necessary truth is intriguing. Pursuing the water example, an understanding of 'water', or the ability to use the term in coherent conversation, does not seem to be dependent on the knowledge that water is H^2O. Water maintains the same *purpose* and defining attributes independent of this knowledge: it is a clear, potable liquid found in rivers and lakes and so on. Adjacent to this point, we could, *prior* to learning that water is H^2O, conceive (imagine, without apparent inconsistency) of water being another chemical compound.[11] Call this compound XYZ and imagine that when lakes and rivers are filled with XYZ, it is indistinguishable from H^2O. From our perspective, knowing that water *is* H^2O, a possible world with XYZ does not have water but something different. However, if we start from a perspective that lacks knowledge of the chemical composition of water, the XYZ world *does* have water. It seems we can conceive of water being XYZ

[10] Braddon-Mitchell, 'Qualia and Analytic Conditionals', 120.

[11] In the final section of this paper I will investigate the notion of being able to 'conceive' of something in this sense.

if we suspend our knowledge of water being H^2O, but so long as we maintain that water is H^2O, the XYZ world is inconceivable.

Let us now develop this line of thinking by introducing the concept of an *indexical*. It is essential to the concept of qualia that it allows an agent to access the intrinsic nature of their experience, and because of this, if an agent believes they have qualia, they cannot be mistaken.[12] It may be useful to compare the statement "I have conscious experience" to the statement "I am here". The definition of "here" is dependent on the subject speaking the utterance; thus I cannot produce an incorrect proposition when I say "I am here", provided that by "here" I mean where I am situated. Similarly, it does not seem possible for a subject to claim they have qualia and make an incorrect proposition. In this sense, qualia is an indexical concept.

Bringing this together, we can now say that if dualism is true in the actual world, and qualia has an essentially non-physical element, then qualia being more than physical facts is an *a posteriori* necessity. Hence, when we conceive of any counterfactual world that lacks non-physical properties, that world does not have qualia. Yet, if the actual world is physicalist, it does not transpire we lack qualia. Because qualia is an indexical, so long as we claim to have qualia, we cannot make an incorrect proposition.[13] From the perspective of physicalist worlds considered as actual, any counterfactual dualist world with non-physical phenomena does not have exclusive access to qualia or even a heightened version of it; they simply have a strange, non-physical additive. So, when dualist worlds are considered as actual, qualia is necessarily non-physical, and when physicalist worlds are considered as actual, qualia is necessarily physical.

Before we proceed, one clarification needs to be made. Concordant with the "epistemic gap" dilemma, philosophers defending dualism do not deny that

---

[12] Braddon-Mitchell, 'Qualia and Analytic Conditionals', 123. Comparatively, I can be mistaken in claiming that I have knowledge of a certain subject matter.

[13] The indexicality here only extends to making the claim about *yourself* having qualia – not other people.

neuroscience provides *markers* of conscious experience, or qualia. For instance, qualia could be accompanied by brain state ∂ (and a causal relationship between ∂ and qualia could even be produced). However, dualists still maintain brain state ∂ does not constitute the *intrinsic nature* of qualia; it is simply a contributing causal factor to something inherently non-physical.[14]

All of this can be practically elucidated with the help of two-dimensional modal semantics, which effectively distinguishes *a priori* and *a posteriori* necessity and their use in language and reasoning. The following will use tables which visually portray these distinctions. The left column are worlds considered as actual, whilst the top row are worlds considered as counterfactual (from the point of view of a world which is considered as actual).[15] We have four worlds: in the first, qualia bears a causal relationship with brain state ∂, but is defined by extra non-physical phenomena. The second world is the same as the first, except the marker for qualia is brain state ß. The third and fourth worlds are physical worlds, where qualia are brain states ∂ and ß respectively.

[14] Braddon-Mitchell, 'Qualia and Analytic Conditionals',117-119.

[15] To use the water example, the actual world is one in which water is H20, however I can counterfactually consider a world in which there is something that looks like water called XYZ.

(N) *"qualia are non-physical states"*

	∂ + non-physical phenomena	ß + non-physical phenomena	∂	ß
∂ + non-physical phenomena	T	T	T	T
ß + non-physical phenomena	T	T	T	T
∂	F	F	F	F
ß	F	F	F	F

We can see that (N) is considered true only when the actual world is one in which qualia is inherently non-physical. For these worlds, (N) holds true even for counterfactual worlds without non-physical properties, as in these worlds, there is no qualia. In physicalist worlds, when we consider a counterfactual world with non-physical states, these are not qualia, but some esoteric additional property.

If we were to consider the statement "qualia are functional brain states", the table for this proposition would be the opposite of the one above. From the point of view of worlds with non-physical qualia, qualia are non-physical states in *all* worlds considered as counterfactual; and from the point of view of worlds with functional brain state qualia, qualia are functional brain states in *all* worlds considered as

counterfactual. Again, we form our conception of qualia based on the world we believe to be actual. This two-dimensional account explains our concept of qualia, whilst also maintaining its malleability subject to what we believe to be actual.

We can now use this framework to account for the zombie intuition. Observe the table for

(Z) *"zombies are possible"*

	∂ + non-physical phenomena	ß + non-physical phenomena	∂	ß
∂ + non-physical phenomena	T	T	T	T
ß + non-physical phenomena	T	T	T	T
∂	F	F	F	F
ß	F	F	F	F

If we draw a diagonal line from the top left to the bottom right, we have the *A-intension* of this statement. The A-intension accounts for our *a priori* intuition, when we have no prior information on what qualia is (whether it is physical or

non-physical) and hence do not know what world is actual. Each of the four horizontal rows are different *C-intensions*, which account for the intuitions we have based on what we think is actual. In this table, the A-intension is contingent, but each C intension is necessary; that is, either necessarily true or necessarily false. A necessary C intension means it is *metaphysically necessary*, given that the actual world is such a way, that the proposition must be true or false. A necessary A intension means that it is an *a priori necessity* that the actual world must be such a way, regardless of a specific detail about the actual world. Tying this back to intuitions, we can imagine that if I am asked if zombies are possible in a ∂ world, and I am certain I am in a ∂ world, then I will confidently respond "no", provided my intuitions are reasonable. However, if I am unsure if I am in a physicalist or dualist world, I will not be able to give such a confident answer. My intuitions will be unclear. If the actual world is a non-physical-qualia world, then zombies are necessarily possible, and worlds ∂ and ß are instances of such worlds. In physical-qualia worlds, zombies are necessarily impossible, and worlds such as *∂ + non-physical phenomena* and *ß + non-physical phenomena* have a foreign additional property that is not qualia.

When we intuit that a philosophical zombie is conceivable, this represents an underlying uncertainty in the truth of physicalism. Even the staunchest physicalist can recognise that if it was discovered tomorrow that non-physical phenomena is intrinsic to qualia, then it would be an *a posteriori* necessity that qualia is non-physical. On the other hand, if we discovered tomorrow that physicalism is true and qualia is ∂ states, then *this* would be an *a posteriori* necessity. Thus, the A-intension represents our concept of qualia with an indeterminate account of what is true actually, whilst the C-intension accounts for our concept of qualia on a pre-set assumption of the way the world is actually. When we conceive of a philosophical zombie it is in virtue of the contingent A-intension. Our ability to conceive of philosophical zombies is in virtue of the credence we give, however small, to dualism being true. The work of determining which world is actual and the (non)existence of

non-physical states can be left to the evidence that neuroscience provides, but regarding the conceivability of P&~Q along the A-intension, this is not a sufficient basis from which to draw ontological conclusions.

3. Chalmers' Response

In his paper "The Two-Dimensional Argument Against Materialism", Chalmers defends the zombie intuition, that is, the assertion that we can conceive of an exact physical copy of ourselves that lacks qualia. He does this by claiming that in the instance of the conceivability argument, it is appropriate to infer ontological conclusions from epistemic premises. Chalmers defines conceivability along the A intension as *primary conceivability* (1-conceivability), which defines what is conceivable on an *a priori* basis. Conceivability along the C intension is *secondary conceivability* (2-conceivability).[16] He concedes that "primary conceivability does not entail metaphysical possibility", which Braddon-Mitchell's argument was instrumental in demonstrating.[17] Metaphysical possibility can (very briefly) be understood as ways things could have been in different possible worlds. Metaphysical possibility and metaphysical necessity is encompassed by logical possibility and logical necessity. To return to our water example, although I can conceive of water not being H^2O along the A-intension, it still has a necessarily false C-intension if the actual world is one in which water is H^2O. There is also a distinction between *prima facie* and *ideal* conceivability, where ideal conceivability is free from human cognitive limitations.[18] When Chalmers makes claims regarding conceivability and possibility, it is understood in terms of *ideal* conceivability. Chalmers posits that although we cannot entail metaphysical possibility from

[16] Chalmers, 'The Two-Dimensional Argument Against Materialism', 317.
[17] Chalmers, 'The Two-Dimensional Argument Against Materialism', 317.
[18] Chalmers, 'The Two-Dimensional Argument Against Materialism', 315. I have concerns about the relevance of ideal conceivability and its relation to possibility, if ideal conceivability is abstracted away from human ability. Arguably, human ability and cognitive constraints are intrinsic to what conceivability is. Unfortunately, these concerns are tangential to this paper.

primary conceivability, we *can* entail "primary possibility" (1-possibility) from ideal primary conceivability.[19] We can utilise the H20/XYZ example to understand how something can 1-possible but not 2-possible. When I "conceive" of a world in which water is not H20 but XYZ, I conceive of a world in which something with identical qualities to water is XYZ instead of H20. Hence, it is 1-possible for water to be something other than H20.

Even if ideal primary conceivability entails primary possibility, the zombie intuition is still firmly separated from metaphysical possibility. In order to bridge the gap between metaphysical and primary possibility, the 1-possibility of P&~Q must entail 2-possibility. This would require the A and C intensions of P&~Q to be the same. In the case of qualia, there is a strong case for the A-intension and C-intension being identical, due to the indexicality of the concept.[20] That is, there is no distinction between the intrinsic nature of qualia and it's appearance/function.

Unfortunately for Chalmers, for any concept about physical objects, this is less convincing. We have already identified that the A and C intensions of water are distinct and from this example it appears that P must have distinct A and C intensions. However, rather than giving more examples about physical objects, Chalmers proceeds by focusing on "microphysical properties". Take, for instance, the physical property of acceleration. Chalmers holds that there could be possible worlds where instead of acceleration being instantiated, there is 'pseudo-acceleration', which fulfills the same function as genuine acceleration but is somehow *not* acceleration. He appears to take the comparison between these microphysical properties and their pseudo counterparts as directly comparable to the case of H^20 and XYZ.[21] As with physical phenomena, Chalmers holds that the A-intension of this property identifies the *role* this property plays, whilst the C-intension identifies what actually plays that

[19] Chalmers, 'The Two-Dimensional Argument Against Materialism', 318.
[20] Chalmers, 'The Two-Dimensional Argument Against Materialism', 320. I.e., compare to the H^20/XYZ case.
[21] Chalmers, 'The Two-Dimensional Argument Against Materialism', 320.

role.[22] It is my stance that this is an inaccurate view to be taken of microphysical properties, and has bizarre ontological implications. However, I will proceed to demonstrate Chalmers' argument before making further criticisms.

For example, take three worlds where 'acceleration' is instantiated by properties A, B and C. If you ride your bike down a hill in property A world, the acceleration, as in property A, is instantiated and makes your bike speed up, and so on with property B and C worlds.

K: *"The bike accelerated"*

	Property A	Property B	Property C
Property A	T	F	F
Property B	F	T	F
Property C	F	F	T

The A and C intensions of K are different. K has a necessary A-intension but contingent C intensions. Additionally, this demonstrates that a world can *verify* (play the right functional role) K without *satisfying* K. To expand, if we analysed the proposition "the bike sped up in an exponential fashion", there would be a T in every square. If we take property B world to be actual, a world satisfies K if acceleration occurs as instantiated in property B. A world merely verifies K if it fulfills the same role but is not the same property. Chalmers proceeds by reasoning that if microphysical properties do not have the same A-intensions and C-intensions, the same must be said for P.

[22] Chalmers, 'The Two-Dimensional Argument Against Materialism', 321.

Let us bring this back to the zombie intuition and recall our present state. P&~Q is conceivable along the A-intension (1-conceivable) because there are some worlds where qualia are non-physical states and can therefore be absent, even if P remains the same. P&~Q is only 2-conceivable under the assumption that dualism is true in the actual world. Having recognised that the conceivability argument cannot be redeemed by directly proving that the A-intension and C-intension of P coincide, he proceeds from this basis to propose to deductively prove that the distinction between the 1-conceivability and 2-conceivability of philosophical zombies entails the falsity of materialism.

Firstly, in virtue of P having different A and C intensions, a world may *verify* P without *satisfying* P. For any statement about a physical property, a world may have this property, even though what actually plays this role is different. Put practically, according to Chalmers, there could be a world that has the property of mass in the sense that it performs all the same functions as mass in the actual world, except instead of these functions being performed by property M, it is performed by property N. Worlds with "pseudo-mass" therefore have the same "structural profile" as the actual world, but possess a different "intrinsic profile".[23] This can be extrapolated to P more generally: there are worlds that appear the same as ours, in the sense that all microphysical properties are fulfilling the same role, but where these microphysical properties are different to the ones instantiated in the actual world. That is, these worlds have the same structural profile, but different intrinsic profiles to our world. Because P has different A and C intensions, so does P&~Q. Therefore, structural properties *alone* do not necessitate the existence of qualia. Chalmers takes these "intrinsic properties" and asks what work they can be doing to necessitate the existence of qualia beyond mere structural properties. Philosophical

[23] Chalmers, 'The Two-Dimensional Argument Against Materialism', 321.

zombies must be missing some intrinsic properties that they do not have and that we, in virtue of our qualia, must have.

Chalmers argues that this proves the falsity of materialism, as the distinction between the A and C intension of microphysical properties necessitates a distinction between their "structural" and "intrinsic" profiles, which thereby demonstrates that structural properties alone do not produce qualia. From this we can deduce that materialism is false and hence we can infer the 2-possibility of P&~Q from 1-possibility, or else that Russellian Monism is true. Very briefly, Russellian Monism posits that there are properties that underlie the structural account of physics, which describes fundamentals and properties only in terms of what they *do,* not what they are. These underlying properties are (wholly or in part) constitutive of qualia. The notion that qualia is inherent in the structural account of physics, and hence physical objects potentially possess some of these underlying properties ("proto-qualia"), means that it is often not accepted as genuine physicalism.

4. Criticism of Chalmers' Argument

4.1 Reliance on Microphysical Properties

I agree with Chalmers that physical phenomena such as water have different A and C intensions. By this I mean that whenever we make a proposition about the intrinsic nature of water (for example, "water is H^20"), this proposition has distinct A and C intensions. What I find problematic is his move from distinguishing the A and C intensions of physical *objects* to what he calls "microphysical *properties*" such as charge, mass and acceleration. Ultimately, I do not think it is possible to distinguish the A and C intensions for microphysical properties. The pressing question is whether this clears up the path for P&~Q having identical A and C intensions, thereby falsifying materialism via an alternate route.

Physical objects supervene on physical fundamentals. Whatever we call "water" in the actual world supervenes on the physical makeup of molecules in the H^20 compound. The supervenience relation between a microphysical property (take mass as an example) and physical fundamentals works a little differently. Microphysical properties are necessarily *properties* held by physical objects and are defined by the roles they perform *in virtue of* the physical makeup of physical objects. For instance, a brick supervenes on physical fundamentals and has a mass of 3kg in virtue of this supervenience relation.

Let us expand on this. Microphysical properties track *relations*, but they are not things in and of themselves. Mass exists *in* the supervenience relation between physical objects and physical fundamentals. Thus, mass tracks a certain identity relation between physical objects and physical fundamentals (the mass of a whole is defined by the amount of physical fundamentals comprising it). Acceleration seems to be more complex, as we have to consider mass and its relation to gravity, which are other microphysical properties. However, I posit that for each microphysical property, we can plot a path that tracks back to a *relation* of some kind that pertains to physical phenomena, whether that be supervenience relation between physical fundamentals and physical objects, or something else (perhaps between physical fundamentals themselves). If we accept that microphysical properties are defined by relations pertaining to physical phenomena, they cannot be 'performed' by a different property.

Building from this point, even if physical fundamentals were entirely different, there is nothing to necessitate that mass would not exist. As long as there is a property we can assign to objects that differentiates the quantity of matter in a physical body, the actual constitution of this matter is irrelevant. Therefore, it is incoherent to say that mass is "performed by property M", and thereby to distinguish distinct A and C intensions. Mass cannot be performed by another property because mass is itself a property.

Thus, I believe the A and C intensions of microphysical properties are necessarily the same. Chalmers does acknowledge this view in his paper, however he does not expand on it beyond a singular mention. To quote Chalmers:

> There are other views of the semantics and metaphysics of microphysical terms that may reject this argument for the distinctness of the primary and secondary intensions of 'mass'. In particular, the argument will not go through on views according to which it is necessary that mass is the property that plays the mass role. … Still, the view sketched above is a quite reasonable view—more plausible than the alternatives, in my opinion— and it is the view that best grounds resistance to an inference from the 1-possibility of P&~Q to its 2-possibility. (Chalmers 2009: 321)[24]

Hence, Chalmers offers no defence of his view of microphysical properties and why it is "more plausible that the alternatives". The problem remains unaddressed by Chalmers: how can the role of a microphysical property be distinct from the property itself?[25]

If the A and C intensions of microphysical properties are the same, this could disrupt the original argument. P can be defined as the conjunction of all the microphysical facts in the actual world.[26] If we were to define a physical fundamental (such as atoms) and provide a complete account for their organisation throughout the universe, then microphysical properties would arise out of that. That is not to say *all* microphysical properties would necessarily arise out of any configuration of fundamentals, because what matters is having the right *relations* between physical

---

[24] Chalmers, 'The Two-Dimensional Argument Against Materialism', 321.

[25] A quick side note: Chalmers' idea that microphysical *properties* have a distinct A and C intension (or intrinsic/structural profile) could display an inherent bias for dualism, or at the very least, a bias against analytic functionalism, as this view separates the property from its functional role. If Chalmers finds it intuitive to make this distinction with microphysical properties, the distinction feels a lot easier with qualia. This is merely an observation and not a critique of the formal argument.

[26] Chalmers, 'The Two-Dimensional Argument Against Materialism', 314. Whilst I gave a broader definition of P earlier as the conjunction of all physical facts, here I am simply specifying microphysical facts as the fundamental facts about fundamentals which comprise the physical facts.

fundamentals (it may be possible to have a world without mass or charge, but this is a separate issue). Perhaps we can say there are a series of ways the world could be constituted and organised so as to produce certain microphysical properties, and this is dependent on there being the right relations pertaining to physical phenomena. In the set of worlds where we have mass, for example, mass is the same property in every one of them. Therefore, on my view the A and C intensions of physical objects are different, but the A and C intensions of microphysical properties that arise out of physical objects are the same. No matter how physical objects are constituted and organised, this will have no impact on the mass, charge and acceleration in a world, provided these properties exist in this world. Comparatively, this is not the case with physical objects. Even if XYZ plays the water role in Twin Earth, it is not water, provided H^20 is water in the actual world.

Here is one way to close this argument in favour of materialism: Provided that P is the conjunction of all microphysical facts, and microphysical properties simply arise out of that, we can maintain that the A and C intension of P are distinct, and hence the A and C intension of P&~Q are distinct. So, the deductive argument does not go ahead. However, it must be granted that no explanation has been given for exactly *how* microphysical properties arise out of physical phenomena. Do they arise through the structural relations, intrinsic relations or both? I have already intimated that they arise out of the structural relations *alone*, because if the intrinsic relations are involved then we have the same problem, that is, that the complete structural account of physics does not account for microphysical properties. But if intrinsic relations have no position in the creation of microphysical properties, what *work* are they doing at all?

Perhaps the apparent redundancy of intrinsic relations produces the same problem with relation to qualia. There could be a world W where P is verified along the A-intension but not satisfied. This world may have XYZ instead of H^20. It also may have 'pseudo' versions of other physical things (perhaps oxygen, atoms, etc.). This world could have the same microphysical properties, such as mass, charge etc. This

world has the same structural relations but different intrinsic relations. Is it necessary that this world has qualia? If it is not necessary, then the structural account of physics alone does not account for qualia, so the 'intrinsic properties' must be doing some work. However, if it is necessary that this world has qualia, then intrinsic relations are again made redundant. Hence, the same problem is recreated through slightly different means, suggesting that the two-dimensional argument against physicalism still bears some force.

4.2 Problematisation of 2-Conceivability

Along with positing a link between 1-possibility and 1-conceivability, Chalmers is also consistent in extending this link to exist between 2-possibility and 2-conceivability.[27]

The conclusion for Chalmers' argument is "materialism is false or Russellian Monism is true".[28] The preceding section criticised the passage taken to this conclusion, however if we discount Russellian Monism for this instance and assume the truth of dualism as the only alternative to materialism, then we must have 2-conceivability, also known as *direct* conceivability of philosophical zombies. From first sight, the idea that we have direct conceivability of philosophical zombies is highly questionable in virtue of the continued existence of this very debate. Whilst Chalmers may claim to have proven that direct conceivability of zombies is possible, the same confidence is not present in our own self-evaluation. Precisely what are we conjuring up when we 'conceive' of a philosophical zombie? I will endeavour to provide some thoughts in response to this question with the aid of Stephen Yablo's work on conceivability. I have chosen to focus on Yablo's work as it is effective in formulating an account of conceivability that escapes the binary between conceivability and

[27] Chalmers, 'The Two-Dimensional Argument Against Materialism', 318.
[28] Chalmers, 'The Two-Dimensional Argument Against Materialism', 322.

inconceivability and thereby produces, in my view, a satisfying account of the subjective experience of 'conceiving' of something that may not be possible.

Stephen Yablo, in his paper "Is Conceivability a Guide to Possibility?" defines the conceivability of proposition p (CON) as being able to "*imagine a world I take to verify p*". Inconceivability (INC) is defined as being unable to "*imagine any world I don't take to falsify p*".[29] Note that upon this account, inconceivability is not simply the negation of conceivability. In fact, the negation of conceivability, ("*I cannot imagine a world that I take to verify p*"), and the negation of inconceivability, (*I can imagine worlds that I don't take to falsify p*), are not mutually exclusive.[30]

Importantly, the account of conceivability as (CON) is in terms of worlds and not just singular situations. In my view, this is linear with our regular experience of conceiving. When I conceive of a proposition, say, (E) "there is a pink elephant waiting for me in my lounge room", I am in one sense imagining a singular situation, however the wider world is implied. I may not imagine all of the details, (for example, the exact shade of pink and the placement of the elephant's feet), but this does not mean I imagine the situation to be abstract in these ways. When I conceive of (E), the elephant I conceive of *does* have a specific shade of pink and exact placement of the feet. Further, if it was not possible to specify these details without avoiding incoherence, then (E) may not be conceivable anymore. In this sense, I imagine the situation *as determinate*, but because I do not spell out every single detail in my conceptualisation, I do not imagine it *determinately*.[31] In the same way, this account necessitates that every proposition that is conceivable entails an entire world, because in theory, it should be possible to detail everything about the world where this proposition is true.

---

[29] Yablo, Stephen (1993) 'Is Conceivability a Guide to Possibility?', *Philosophy and Phenomenological Research* 53: 29.

[30] Yablo, 'Is Conceivability a Guide to Possibility?', 31.

[31] Yablo, 'Is Conceivability a Guide to Possibility?', 28.

Further, under Yablo's account, there is room for the zombie intuition to be undecidable, that is, neither conceivable nor inconceivable. We can potentially have the conjunction of these two statements, which are the negation of (CON) and (INC):

(~CON) I cannot imagine any world in which P&~Q is verified.

(~INC) I can imagine worlds in which it is not falsified that P&~Q.

(~CON & ~INC) I cannot imagine any world in which P&~Q is verified, and I can imagine worlds in which it is not falsified that P&~Q.

The conjunction of these propositions could entail there are worlds where P&~Q is in between true and false, or else its truth value is ambiguous. That is:

(a) I can imagine a world in which P&~Q has a truth value that is in between true and false.

OR

(b) I can imagine a world in which P&~Q is not verified, but it is also not falsified, because the situation is ambiguous.

In order to avoid a discussion on non-classical logic, we shall proceed with interpretation (b) of undecidability.

This interpretation of undecidability can also be elucidated in terms of two-dimensional semantics and Chalmers' primary and secondary conceivability. Suppose that a proposition is 1-conceivable if there is at least one T in its A-intension. Suppose further that the criterion for 2-conceivability of a proposition is that there is at least one T in a given C-intension.[32] Propositions that are conceivable according to definition (CON) align with 2-conceivablility but not 1-conceivability, as (CON) entails being able to imagine the situation *determinately*, which 1-conceivability in isolation discounts. Further, propositions that are undecidable according to (b) are

[32] If there is only one T in a certain C intension, this should occur when the actual and counterfactual world align (granted that it is impossible to have 2-conceivability without 1-conceivability).

1-conceivable but not 2-conceivable. In these cases, we can roughly conceive of a situation where a proposition is verified (that is, along the A-intension), but we cannot, even in principle imagine the situation determinately, because we do not know what the actual world is like. Propositions that are inconceivable according to (INC) are neither 1-conceivable nor 2-conceivable.

4.3 Assessment of "Zombies Are Conceivable"

When I imagine a philosophical zombie, do I imagine it *determinately*? If pressed, could I provide all the details about a philosophical zombie? If qualia is non-physical, all we would have to do is imagine the affirmation of being a physical duplicate and the negation of the non-physical phenomena constituting conscious experience, and provide every singular granular detail. That being said, providing every single detail is an impossible bar to set.[33] However, if we believe qualia is non-physical in the actual world, we are able to *in principle* imagine a philosophical zombie determinately and thus have 2-conceivability. Yet if qualia is physical, we cannot obtain 2-conceivability even in principle. Put simply, the 2-conceivability of qualia rests on the assumption that it is non-physical.

The 'conceivability' of zombies along the A intension betrays our lack of certainty in the physicalist thesis. If we discovered beyond doubt that the world is physicalist, then zombies would be inconceivable, as we would not view dualist worlds as genuine candidates for the actual world. Our *a priori*, or primary conceivability of

[33] There are interesting arguments to be made regarding what details are relevant when conceiving of a situation. If the bar is set too low, critical details (such as if qualia is physical or non-physical) could be missed out, whereas if the bar is set too high, the criterion could end up being to detail an entire possible world. The latter seems to me a better bar to set if a correlation between conceivability and possibility is to be posited. Surely if all the relevant details about a situation can be articulated without inconsistency, then it is metaphysically possible. However, there may be seemingly unconnected details that, once articulated, actually produce a contradiction. Perhaps it is only once an entire possible world is articulated that we can be sure of the metaphysical possibility of a 'conceived' state of affairs. Yet at this point, we are positing something that looks more like an ersatz theory of possible worlds rather than conceivability.

zombies can be clarified with Yablo's account of undecidability, interpreted according to (b). When we try to imagine a philosophical zombie from a basis of ignorance about which world is actual, what we imagine is by nature indeterminate. We cannot, even in principle, spell out the details of the world in which P&~Q is true, because we don't know what actual world our conceived situation is based on and therefore what we are negating in terms of Q. Therefore, our 1-conceivability of philosophical zombies can be explained by Braddon-Mitchell's account, that is, we are not one hundred percent certain which world is actual and therefore posit conceivability claims on this shaky foundation. Our purported 2-conceivability of zombies only holds on the assumption that dualism is true, which is a premise the conceivability argument aims to prove.

Chalmers' argument purports to prove that either materialism must be false, giving us 2-conceivability of zombies, or else Russellian Monism is true. However, our direct experience only gives us 1-conceivability, and 2-conceivability seems relegated to those who already have a firm dualist intuition.

5. Conclusion

We can now return to Braddon-Mitchell's original account for the zombie intuition in terms of two-dimensional semantics. In one sense, all conceivability and possibility claims are founded on a conception of what the world is actually like, meaning that P&~Q is conceivable for dualists and inconceivable for physicalists. Nevertheless, physicalists can acknowledge they may be mistaken, and thereby P&~Q is conceivable along the A-intension, in light of underlying uncertainty about what the actual world is like. Chalmers responds by accepting that conceivability along the A-intension does not entail metaphysical possibility, but that conceivability along the C-intension does, and further it can be deductively proven that the A and C intensions of P&~Q are identical. This argument at least partially relies on a sleight of hand that shifts from discussing physical objects with intrinsic profiles that differ to

their structural profiles, to "microphysical properties" with differing intrinsic and structural profiles. The latter is incoherent. However, if we posit that microphysical properties have identical A and C intensions, then the same issue arises. That is, if structural relations alone do not necessitate qualia, then what are these intrinsic relations and what further work can they be doing?

If Chalmers' argument against materialism is accepted, this implies that we have secondary, or direct conceivability of zombies. This notion can be investigated with recourse to Yablo's work on conceivability and inconceivability according to (CON) and (INC). (CON) pertains to direct conceivability, and direct experience suggests we do not have direct conceivability of philosophical zombies, unless we already have a firm predilection to dualism. In the absence of direct conceivability, the primary conceivability of philosophical zombies remains intact, and can be interpreted according to Yablo's account as indeterminate, sitting between conceivability and inconceivability due to ambiguity surrounding the state of the actual world.

References

Braddon-Mitchell, David (2003) 'Qualia and Analytical Conditionals', *The Journal of Philosophy* **100**: 111–35. https://doi.org/10.5840/jphil2003100321.

Chalmers, David (2009) 'The Two-Dimensional Argument Against Materialism', in A. Beckermann, B.P. McLaughlin, and S. Walter, eds. *The Oxford Handbook of Philosophy of Mind*: 313-336. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199262618.003.0019.

Jackson, Frank (1986) 'What Mary Didn't Know', *The Journal of Philosophy* **83**: 291-95. https://doi.org/10.2307/2026143.

Yablo, Stephen (1993) 'Is Conceivability a Guide to Possibility?', *Philosophy and Phenomenological Research* **53**: 1–42. https://doi.org/10.2307/2108052.