Is Mereology a Guide to Conceivability?
[forthcoming in Mind]

Daniel Giberman
dgiberman@gmail.com

Zombies are unconscious objects with conscious physical micro-duplicates. If zombies are possible then physicalism (the thesis that the physical determines the mental) is false. It has been argued that zombies are possible if conceivable for an agent with ideal rationality. At any rate, they are possible only if so conceivable. This essay uses a mereological constraint to highlight the fine-grained differences between actually conscious physical objects and certain of their actually consciousness-incapable proper parts. These mereological considerations form the basis of an argument by dilemma that zombies are inconceivable. Either an arbitrary actually conscious object might have had simpler consciousness-capable parts (and more complex consciousness-incapable parts) than it in fact has, or not. The affirmative horn leads to a version of panpsychism that is inconsistent with the ideal conceivability of zombies. The negative horn rules out zombies as incoherent. The upshot is a new reason to deny the conceivability of zombies.

1. Introduction

Suppose I have a microphysical twin: his most basic physical parts have the same intrinsic properties and stand in the same spatiotemporal relations as mine. But, while he is typing on his computer and blinking his eyes just as I am, his mental states wholly lack phenomenal character. There is nothing it is like for him to hear the clicking keys or see the text on his screen. He is a zombie. If he is a possible object then the thesis that, necessarily, the physical facts determine the mental facts (hereafter ‘physicalism’) is false. This kind of thought experiment probably has something in common with the back of your hand. You are also likely to be familiar with complaints about the claim that conceivability secures possibility, upon (a certain version of) which the zombie case turns. (If not, see Yablo 1993, Gendler and Hawthorne 2002, and Chalmers 2010a). A less familiar though not unprecedented complaint is that zombies are not even conceivable (Van Gulick 1993, Dennett 1995, Kirk 1999 and 2008, Botterell 2001, Worley 2003, Marcus 2004). There are numerous ways to argue for inconceivability. One way is to claim that if we had powers of conception that reflected idealized rational ability and empirical knowledge of physical facts at the actual world then
we would find zombies conceptually self-defeating. Call this version of the inconceivability objection to zombies ‘ideal inconceivability’. Friends of zombies have replied that there is no plausible way to make ideal inconceivability sufficiently precise (Chalmers 2002a, 2010a, 2012). Whatever information an ideal conceiver may have about physical microstructure, it will only ever entail information about structures that are derivable from physics; but consciousness facts, the friends of zombies say, are not facts about structures derivable from physics; so, for arbitrary physical microstructure X, not even an ideal agent will be able to rule out as inconceivable all states of affairs in which an unconscious object instantiates X.

The present essay provides a new defence of ideal inconceivability by articulating a mereological constraint on actual bearers of phenomenal consciousness.¹ According to the constraint, paradigmatic actually conscious objects are:

(i) mereologically complex, and
(ii) capable of losing some parts while retaining consciousness.

It will be argued that properly respecting this constraint either renders zombies incoherent or leads to a zombie-undermining view of the relationship between consciousness and basic physical properties.

The qualification ‘paradigmatic’ and the ascription of ‘consciousness’ as employed in the constraint require clarification. Paradigmatic actually conscious objects are mereologically complex macro objects that are ‘maximal’ in the loose but intuitive sense that they are not undetached proper parts of ordinary objects. Examples include human beings and other conscious animals. Examples of objects outside the class of paradigmatic conscious objects include undetached proper parts of human beings. On present usage, ‘conscious’ picks out an

¹ Unless otherwise noted, I will use ‘consciousness’ as short for ‘phenomenal consciousness’.
intrinsic phenomenal property that may be exemplified equally by many of a given conscious object’s proper parts, though perhaps only upon detachment. So, on present usage, the claim that some mereologically basic physical object $x$—a quark, say—is conscious is the claim that $x$ has the same general property of being a subject of phenomenal experience that you and I enjoy. It is not the claim that $x$ merely has some proto-phenomenal property that would help ground full-blown phenomenal experience for some complex whole in the presence of an appropriate arrangement of other basic objects; nor is it the claim that $x$’s physical properties are ultimately grounded in some sort of idealism, though it is consistent with this latter claim.  

Let us use ‘quark’ as an abbreviation for ‘basic element in a spatiotemporal-cum-mereological physical structure’. And let us assume for the time being that no extracranial or particularly small proper parts of paradigmatic conscious objects like arms, cells, and quarks are actually conscious. My arms, cells, and individual quarks notwithstanding, I have many proper parts that either are conscious or would be if they were to be detached, for example, *me-minus-an-arm*. Call the disjunctive property of being either conscious simpliciter or conscious-if-detached ‘consciousness-capability’. A question of central importance in what follows can now be stated: *exactly which objects outside the class of paradigmatic conscious objects are consciousness-capable, both actually and possibly?*

---

2 See the discussion of ‘Type-F Monism’/‘panprotopsychism’ in (Chalmers 2002a).

3 The difference between consciousness simpliciter and consciousness-capability is important for undetached proper parts whose putative conscious states would be very similar to those of the relevant whole object. For humans these are relatively large, brain-inclusive proper parts. For example *me-minus-an-arm*, if conscious while undetached, has conscious states that are extremely similar to my own. This gives rise to a puzzle about overpopulation of objects in my vicinity that have identical or near-identical conscious states. The distinction between consciousness simpliciter and consciousness-capability circumvents such puzzles. Since there is no such puzzle for small, brain-exclusive proper parts (for example, quarks), the distinction between consciousness simpliciter and consciousness-capability may be allowed to blur in discussion of such objects. Moreover, it is far less clear what the criteria for being detached or undetached are in the case of increasingly low-level proper parts like quarks than it is for high-level parts like arms. This is another reason to let the distinction blur in the low-level cases. Accordingly, consciousness-capability for quarks and similarly low-level parts will be understood as consciousness simpliciter in the sequel.
For any conscious creature $x$ at arbitrary time $t$ from whose microphysical structure we are asked to conceive a zombie, there will be a fine-grained mereological spectrum from the (conscious) whole of $x$ at $t$ to its (presumably consciousness-incapable) most basic physical proper part(s) at $t$.\(^4\) Take my case as an example. At one pole of the spectrum (the ‘whole pole’) will be me and at the other (the ‘simple pole’) will be my most basic physical spatiotemporal-cum-mereological unit, a ‘quark’ in present terminology. In between will be all my undetached proper parts. This is an expansive and varied lot. It includes the bearers of such descriptions as ‘me-minus-a-quark’, ‘me-minus-a-neuron’, ‘me-minus-an-arm’, ‘a fusion of the easternmost half of my left ear and three cells in my right big toe’, ‘a fusion of a quark in my brain and a quark in my arm’, and the like. Points on the spectrum correspond to quark cardinality. So the point on the spectrum adjacent to the whole pole contains all my parts that fit the description ‘me-minus-a-quark’; and the point adjacent to the simple pole contains all my parts that are two-quark fusions.

The force of the mereological constraint is that it entails—still assuming that basic proper parts are never actually conscious—that somewhere along the mereological spectrum of a paradigmatic conscious object is a point containing that object’s \textit{mereologically minimal consciousness-capable undetached proper parts}. This is the collection of parts every member $x$ of which is such that:

(i) if $x$ were to become detached at the actual world, it would be conscious, and

(ii) if $x$ then lost even a single quark (without subsequent replacement by another), it would fail to be conscious.

\(^4\) For simplicity, the relativization of microphysical structures to times will be left implicit hereafter.
There must also be a point containing some *mereologically maximal consciousness-incapable undetached proper parts*. This is the collection of parts every member $x$ of which is such that:

(i) if $x$ were to become detached at the actual world it would then need to gain as a part only one additional quark (properly placed) in order to be conscious, and

(ii) no parts that satisfy the condition described in (i) are more complex than $x$.

Call the segment on a given object’s mereological spectrum that is bounded by these two points its ‘mereological threshold for consciousness’ (MTC) since all the object’s parts that lie beyond one end of the threshold are consciousness-capable and all the parts that lie beyond the other end are consciousness-incapable.

An ideal conceiver with ideal knowledge of actual physical facts would know exactly which proper parts of me are to be found within, without, and at the boundaries of my MTC. The ideal conceiver knows all physical facts because she must be able to fill in arbitrary details to check for ideal coherence when imagining a physical duplicate of the actual world. Among these facts are, for example, facts about exactly which portions of my brain could be removed at the actual world before I no longer exhibit outward signs of being conscious. That these outward signs (whether functional, behavioral, or causal) indicate consciousness is supported by the required assumption that I am not a zombie. Granted, facts about some such portions of my brain are currently available to non-ideal agents like us, but the facts available to us are much more coarse-grained than the facts available to the ideal conceiver. The importance of this difference in fineness of grain will be emphasized below.

Notice that the knowledge here attributed to the ideal conceiver is not ideal knowledge of what is *metaphysically* possible for me—which would threaten to trivialize the dialectic of ideal conceivability—but merely ideal knowledge of exactly which of my parts are *actually*
consciousness-capable (on the assumption that I am not a zombie). Indeed, an ideal conceiver would know, with respect to any conscious creature \(x\) at the actual world, exactly which proper parts of \(x\) are actually consciousness-capable. That is, the ideal conceiver would know down to the finest grain in the spatiotemporal-cum-mereological spectrum which physical structures are sufficient for consciousness at the actual world.

Once it is granted that an ideal conceiver would have knowledge of fine-grained differences between consciousness-capable and consciousness-incapable physical spatiotemporal-cum-mereological structures at the actual world, clarity is gained on two fronts. First, reasons beyond mere appeal to vaguely described ideal knowledge are furnished for thinking that facts about consciousness might well be inferable from facts about physical structure. There is great variation across actually consciousness-capable objects. Yet, for any actually consciousness-capable object \(x\), there is little variation between an arbitrary mereologically minimal consciousness-capable undetached proper part of \(x\)—call it ‘\(y\)’—and an arbitrary mereologically maximal consciousness-incapable undetached proper part of \(y\). So an ideal conceiver would have knowledge of links between physical structure and consciousness that are both applicable across a great variety of structures and highly sensitive to fine-grained structural differences. Such breadth and sensitivity of knowledge about how structure and consciousness actually relate is evidence (albeit inconclusive) of an ability to derive consciousness facts from structure facts across states of affairs that are physically similar to the actual world, including non-actual states of affairs that a lesser conceiver might believe to contain zombies.

Second, resources are furnished for explaining why we lesser conceivers are tempted to think that the inconceivable is conceivable. Since our powers of conception and our
knowledge of physical structures are comparatively coarse-grained, we are susceptible to mistaking a mereologically maximal consciousness-incapable proper part of something for a consciousness-capable part when we attempt to imagine zombies. By analogy, suppose Smith is ‘eyeballing’ the doorway of a prospective apartment, attempting to conceive that his favourite sofa could fit through it. Unfortunately for Smith, the sofa is prohibitively wide by a miniscule margin. An ideal conceiver would know the measurements of the doorway and the sofa and consequently would rule out the sofa’s fitting as inconceivable, just as she would be able to discern me from my mereologically maximal consciousness-incapable proper parts. Smith, in contrast, mistakes the doorway in question for a sufficiently wide one in his attempt to conceive of the sofa’s fitting through, for he lacks knowledge of the relevant measurements. Similarly, we misidentify facts about my mereologically maximal consciousness-incapable undetached proper parts—such as the fact that they are consciousness-incapable and the fact that they exist detached in many states of affairs that are physically near-indiscernible from the actual world—as being facts about a zombie.

In the next section, a new argument from the above mereological constraint to the ideal inconceivability of zombies is presented. The argument proposes a dilemma based on the tautology that either an object’s MTC is contingently located on its mereological spectrum or it is not. If it is then, for reasons to be explained, it is highly plausible that a version of panpsychism that necessitates consciousness for physical structures is possible. Yet if MTC location is not contingent then physical duplicates of actually conscious objects must be conscious, for they lie on the uniformly consciousness-capable ends of the relevant mereological spectra. Either way, zombies come out inconceivable. Section 3 discusses a typical zombie thought experiment and diagnoses where the defender of zombies errs in
taking it to be ideally conceivable. The general diagnosis is intimated in the above sofa example: we are prone to mistaking certain consciousness-incapable proper parts of conscious objects for the objects themselves. Section 4 considers and responds to objections, including the arguments against ideal unconceivability presented in (Chalmers 2002a).

2. Parting with zombies

The most basic zombie argument goes as follows. If zombies are metaphysically possible then physicalism is false; zombies are metaphysically possible; so physicalism is false. The second premiss is thought to follow from the conceivability of zombies and the thesis that conceivability entails metaphysical possibility. The relevant notion of conceivability in the most interesting version of the basic argument is ideal, positive conceivability: conceivability that respects imaginability and internal coherence upon ideal rational reflection (Chalmers 2002b).

The more sophisticated zombie argument (Chalmers 2010a) invokes a two-dimensional modal framework, which distinguishes primary from secondary conceivability and possibility, respectively. (It is primarily conceivable that water is not H₂O because it is coherently imaginable that something other than H₂O plays the conceptual ‘water role’ of being a clear, flavourless liquid that constitutes rivers, etc. But it is not secondarily conceivable given that H₂O uniquely plays this role at the actual world. A proposition or state of affairs is primarily possible if there is a world at which it is true or obtains, as primarily conceived.) The sophisticated zombie argument also has a more moderate, disjunctive conclusion that takes into account the nuance of recent attempts to characterize physicality.

---

5 Unless otherwise noted, ‘conceivability’ hereafter picks out ideal primary positive conceivability.
Here is the sophisticated version. Zombies are ideally, primarily conceivable; if zombies are ideally, primarily conceivable then they are primarily possible; if they are primarily possible then either they are secondarily possible or Russellian monism (roughly, the thesis that consciousness and physical dispositions are both underwritten by the same intrinsic properties of basic objects) is true; if they are secondarily possible then physicalism is false; so either physicalism is false or Russellian monism is true (Chalmers 2010a, p. 152). This disjunctive conclusion is a serious one for physicalism in so far as Russellian monism—which entails that we lack direct empirical access to the intrinsic properties that underwrite physics—deviates from empiricism in a way that is uncomfortable to many physicalists.

The sophisticated zombie argument is valid. What is controversial, of course, is whether it is sound. The first premiss is the target of the present essay. The second premiss and its kin have been the target of many others. The role of the second premiss for present purposes will be discussed in the next paragraph. The thought in favour of the third premiss is that the most plausible way to separate the primary possibility of zombies from the secondary possibility is to trade—as Russellian monism does—on the distinction between the theoretical roles of physical properties and the players of those roles at the actual world. To the extent that it is plausible that the players play their roles only contingently, primary and secondary possibilities involving physical indiscernibility come apart. Finally, given certain plausible assumptions about the modal nature of physicalism, the fourth premise follows without much controversy from the concept of a zombie.

Much like the simple version of the zombie argument, the sophisticated version invokes a conceivability-possibility linking sub-premiss as the basis for its second premiss:
(CP) Ideal, primary conceivability entails primary possibility.

It will be argued presently that (CP) and some plausible observations about the mereological constraint articulated in section 1 jointly entail that zombies are not primarily conceivable. So the sophisticated argument is unsound.

To begin to set the stage for the anti-zombie argument, we should note that the ideal primary conceivability of zombies requires that what I will call ‘physical panpsychism’ is primarily impossible. Physical panpsychism is the thesis that phenomenal consciousness is an intrinsic categorical property of mereologically basic particulars, which property plays a constitutive, underwriting role in (i) the fundamental properties of ‘final’ physics at the actual world and (ii) the exemplification of consciousness by more complex structures. To see that the primary conceivability of zombies is inconsistent with the primary possibility of physical panpsychism, notice that physical panpsychism has actuality built in: it is a thesis about actual final physics. So even the primary possibility of physical panpsychism would entail that actual physics presupposes consciousness. Consequently, one cannot coherently conceive of a state of affairs that is physically indiscernible from the actual world—as required by the primary conceivability of zombies—unless either the physical structures in that state of affairs are conscious or physical panpsychism is assumed primarily impossible. This consequence holds even if the state of affairs in question is, per primary conceivability, not assumed to match a posteriori facts about the actual world, for the very content of physical panpsychism—if not rejected as impossible—entails facts about the deep nature of the physical at the actual world.
Continuing with the stage setting, we should recall that the mereological constraint articulated in section 1 implies the existence of a mereological threshold for consciousness (MTC). If we set panpsychism aside, the lower bound of my MTC probably includes some middle-sized proper parts of my brain; the lower bound of an octopus’s MTC probably includes some middle-sized proper parts of its brain; and traditionally unconscious objects like mountains and popsicles do not have MTC’s at all. The question of central importance in the sequel is whether a given object’s MTC is only contingently located on that object’s mereological spectrum (assuming its mereological spectrum itself to be held fixed across worlds). Put differently: could a physical twin of an actually conscious object have simpler consciousness-capable proper parts or more-complex consciousness-incapable proper parts than the object actually has? The argument will be that this question introduces a detrimental dilemma for the proponent of zombies. If the answer is negative then zombies are inconceivable for an ideal conceiver. Yet if the answer is affirmative then physical panpsychism is possible and, for reasons described in the preceding paragraph, zombies again are ideally inconceivable. Here, finally, is the argument in more detail.

1. If zombies are ideally primarily positively conceivable then physical panpsychism is not primarily possible. (See two paragraphs up.)

2.1 For arbitrary actually conscious physical structure \(x\), either it is ideally primarily positively conceivable that \(x\)’s MTC could have been different or it is not. (Tautology)
Suppose it is not. Then $x$’s MTC is guaranteed to have an upper bound (which is less than $x$ itself) in every conceivable state of affairs in which $x$ exists. So, since every point on $x$’s mereological spectrum beyond the MTC’s upper bound contains consciousness-capable parts of $x$, it is inconceivable for $x$ itself not to be conscious. Since $x$ is an arbitrary actually conscious physical structure, it follows that zombies are inconceivable. For the sake of continuing the argument, then, the present supposition is to be rejected. So:

2.2 It is ideally primarily positively conceivable that $x$’s MTC could have been different.

3. If it is ideally primarily positively conceivable that $x$’s MTC could have been different then it is ideally primarily positively conceivable that physical panpsychism is true. (This premiss is defended below.)

4. If it is ideally primarily positively conceivable that physical panpsychism is true then physical panpsychism is primarily possible. (CP)

5. Physical panpsychism is primarily possible. (2.2, 3, 4, modus ponens) So:

6. Zombies are not ideally primarily positively conceivable. (1, 5, modus tollens)

Only premiss 3 requires further argument (though one might also object to the idea of a precisely bounded MTC, on which more in section 3 below).
Here is the argument for premiss 3. Within the present dialectic, there are no criteria other than imaginability and expected coherence upon ideal rational reflection for determining where a given object’s MTC could conceivably fall on its mereological spectrum, given that it could conceivably fall differently than it actually does. And from the non-committal perspective of the antecedent of premiss 3, all physical spatiotemporal-cum-mereological structures are equal with respect to being imaginably and coherently consciousness-capable. So all physical spatiotemporal-cum-mereological structures are equal with respect to whether we presently have reason to believe that an ideal agent would conceive that they are consciousness-capable. It follows that it is conceivable that the lower bound of an object’s MTC could fall anywhere on its mereological spectrum, including the simple pole. Yet once it is granted as conceivable that mereologically basic physical items at the simple pole exemplify phenomenal consciousness, there is no bar to the conceivability of the rest of physical panpsychism. There is no apparent reason to deny that an ideal conceiver would allow as coherently imaginable that the basic items’ consciousness (i) underwrites the items’ fundamental physical properties and (ii) gives rise to the consciousness of more complex objects like brains.

One might worry that this argument for premiss 3 is too quick. If we consider more nuanced conceptions of what it is to be a fundamental physical property, the worry goes, then reasons emerge for thinking that an ideal conceiver might well reject the connection between conscious basic objects and the underwriting of fundamental physical properties. Following (Stoljar 2001), we should distinguish between a conception of physical theories as committed to both categorical and dispositional properties, on the one hand, and a conception of physical

---

6 See (Brown 2010) for a related defense of the parity between the roles played by conceivability in pro- and anti-zombie arguments, respectively.
theories as committed only to dispositional properties, on the other. The former conception, which is concerned in part with categorical properties of objects, is the ‘o-conception’ of the physical. The latter, concerned only with the dispositional structure tracked by physical theories, is the ‘t-conception’.7 The worry for premiss 3 is that the exemplification of categorical consciousness by spatiotemporal-cum-mereologically basic objects may well be insufficient for securing the dispositional properties that, according to the t-conception of the physical, comprise fundamental physics. This is because, on certain plausible ways of understanding the nature of dispositional properties, no specific categorical properties are such that their exemplification is sufficient for the exemplification of dispositional properties by a basic object in a given state of affairs (even if dispositional properties must be grounded in some or other categorical properties). So the basic objects’ consciousness, which is a categorical property, cannot ‘underwrite’ or be ‘constitutive of’ dispositions in the sense required by the t-conceptual view of physical panpsychism since no categorical properties whatsoever are equipped for that task. Rather, the dispositions may be held fixed independently of the intrinsic nature of the categorical bases. Accordingly, if the t-conception is correct and if disposition distributions are independent of categorical distributions in the way described, then an ideal conceiver would reject the leap from the exemplification of categorical consciousness by basic objects to any associated underwriting of fundamental physical properties, rendering physical panpsychism less than ideally coherent.

One response to this worry for the coherence of physical panpsychism is to motivate the o-conception of the physical, for the o-conception allows a place for categorical

---

7 Strictly speaking, one’s version of the t-conception turns on one’s particular physical theory. Accordingly, the t-conception need not exclude categorical properties, for one might take physics to range over categorical properties. However, since categorical-friendly t-conception is uninteresting for present purposes, I simplify exposition by building commitment to exclusively dispositional physical theories into the t-conception itself.
consciousness among the properties responsible for the physical character of basic objects; and this allowance is independent of concerns about dispositions. One motivating factor for the o-conception is that it meets a plausible criterion of adequacy on any conception of the physical, namely, that it must cover all properties exemplified in spacetime, even if it does not thereby rule out extra-spatiotemporal entities. The t-conception, by contrast, leaves open the possibility that some properties exemplified at spatiotemporal locations are extra-physical. So, this first response to the worry for physical panpsychism concludes, the o-conception is superior to the t-conception. While this response carries some force, it will be worthwhile to consider a second response that treats the t-conception on its own terms, for theorists friendly toward or neutral about zombies may hold the t-conception for independent reasons.

The second response to the worry that the t-conception threatens the coherence of physical panpsychism requires an interlude on the metaphysics of properties. The response is that the worry turns on an imperspicuous ontology of property exemplification that overlooks two distinctions. The first distinction is between properties (which have been understood variously to be universals, classes of tropes, sets of objects, meanings of predicates…) and property exemplifications (which have been understood variously to be universals-at-locations, universal-bare particular pairings, tropes, states of affairs…). My sweater and the sky both exemplify the property blueness, but they do so with distinct exemplifications of

---

8 See (Stoljar 2001) for more points in favor of the o-conception.

9 Whether understood as tropes, universals-at-locations, or universal-bare-particular pairings, property exemplifications are spatiotemporally located (see Campbell 1981 and 1990, Simons 1994, Armstrong 1978, 1988, 1989, 2004, and 2010). The chief difference between trope theory and universal realism concerns numerical identity: the distribution of a multiply-exemplified qualitative property in terms of trope theory involves more than one trope; in terms of universals it involves just one universal but more than one location. There are myriad theories of how ‘substantial’ objects like tomatoes relate to tropes or universals, respectively (see my 2012), but the details are unimportant for present purposes. As is about to be made clear in the text, the important point is that—however a given ontology of properties describes property exemplifications—the t-conception does best to treat the claim that the basic physical items are conscious as saying that spatiotemporal-cum-mereologically basic disposition exemplifications are conscious.
blueness. One way to recognize that the exemplifications are distinct is to recognize that the sky’s blueness exemplification takes up much more space than does the sweater’s. The second distinction is between a *disposition exemplification* and a *categorical basis for a disposition*.10

The worry about the coherence of physical panpsychism highlights the fact that, on the t-conception, the physical properties—the dispositions—do not depend on specific categorical properties (even if they contingently are grounded by some or other categorical base).

However, the most basic *physical* elements in the world’s spatiotemporal-cum-mereological structure, on the t-conception, are just really spatially small disposition *exemplifications*, which themselves are neither traditional objects or substances, nor categorical bases (thus talk of ‘items’ in the argument for premiss 3). Rather, they are property exemplifications in a sense that is neutral with respect to debates in the ontology of properties over tropes, universals, bundle theory, and bare particulars. Through the lens of the t-conception, the most perspicuous way to understand consciousness-capability for basic physical elements in some creature’s spatiotemporal-cum-mereological structure is as entailing that a given disposition exemplification—whether a concrete trope or a universal-exemplified-at-a-location or a universal-paired-with-a-bare-particular—is itself conscious.11

For example, consider some maximally specific charge quantity F as understood in contemporary physics. Now consider two metaphysically possible mereologically basic exemplifications of F, F’ and F”, where F’ is conscious (there is something it is like to be that very disposition exemplification) and F” is not. Are F’ and F” physically equivalent? Yes and

---

10 It is important to recognize that disposition exemplifications are not the same as disposition *manifestations*. Unmanifested dispositions still have exemplifications.

11 Notice that this approach does not render consciousness a polyadic relation. Properties of disposition exemplifications need not be polyadic.
no. Yes in the sense that both bear the dispositional profile of the relevant charge property. But no in the sense that F—despite being a maximally specific charge quantity—turns out (given the example’s stipulation that \(F'\) and \(F''\) are both possible) to have two kinds of exemplifications when distributed across metaphysically possible worlds: it can either have conscious exemplifications or unconscious exemplifications. Thus, on the assumption that a complete physics will account for all possible manner of physical property exemplifications, conscious and unconscious F exemplifications are distinct physical elements. Abstracting now from charge F to consider any fundamental physical disposition \(\Phi\), there will be conscious and unconscious exemplifications of \(\Phi\) distributed across metaphysically possible worlds. The physical panpsychist claim that all the basic physical elements are conscious (which, through the lens of the t-conception, amounts to the claim that all the basic disposition exemplifications are conscious) entails that, for all dispositions \(\Phi\) involved in the basic physical elements, only the conscious exemplifications of \(\Phi\) obtain. Accordingly, each \(\Phi\) exemplification will exemplify the phenomenal property what it is like to be a \(\Phi\) exemplification. Moreover, the phenomenal property what it is like to be a \(\Phi\) exemplification cannot characterize \(\Psi\) exemplifications, for any property \(\Psi \neq \Phi\); it can only characterize \(\Phi\) exemplifications. So the distribution of the categorical phenomenal property what it is like to be a \(\Phi\) exemplification is sufficient to fix the distribution of the disposition \(\Phi\). It follows that the ascription of consciousness to all the basic physical elements of arbitrary world \(w\) determines the distribution of all the fundamental physical dispositions at \(w\), even if those dispositions are not fixed by any categorical base exemplified by the substantial object in question. This consequence vindicates the coherence of physical panpsychism, even while
viewed through the lens of the t-conception and even while suspending endorsement of the claim that an object’s categorical properties fix its dispositional profile.\textsuperscript{12}

The interlude on dispositions is not quite complete, however. Some ontologists may reject any numerical distinction between disposition exemplifications and categorical bases. For example, David Armstrong (1996) holds that dispositions are contingently identical to their categorical bases: disposition F in this world just is its base G (a categorical universal or collection of categorical universals); but in other worlds it just is base H (≠G). One problem for this view is that it contravenes the rigidity of property names (see Choi and Fara 2012, who reference Prior et al. 1982). But I think the view as stated is not quite what the Armstrongian should say when it comes to exemplifications of dispositions. The more perspicuous thought is that disposition exemplifications are not categorical exemplifications simpliciter, but rather categorical exemplifications playing some causal role with respect to other categorical exemplifications. In Armstrong’s ontology, this is to say that disposition exemplifications are really exemplifications of higher order Armstrongian states of affairs: exemplified relations between (or among) first order categorical universals.\textsuperscript{13} The relatum that causes the exemplification of the other relata is the categorical base of the disposition; but the disposition exemplification requires the exemplifications of the causally efficacious relatum and all the relevantly caused relata (or relata that would-be-caused-upon-manifestation). So through the t-conceptual lens of the Armstrongian view, the most basic disposition

\textsuperscript{12} This second response does not render the t-conception inconsistent with the conceivability of zombies, for it may be that only highly mereologically complex exemplifications of dispositions are ever conscious and that these are conscious only contingently. What the response commits the t-conception to is the claim that if the most mereologically basic physical items are conscious then mereologically basic disposition exemplifications are conscious.

\textsuperscript{13} Armstrong (1996 and 2010, p. 50) approaches the spirit of this suggestion when he emphasizes that manifested dispositions are explained by laws of nature, which in his ontology are just a special, causally relevant subset of second-order relations between/among universals.
exemplifications will involve not just one categorical relatum but rather exemplifications of
the second order relation of causation between/among two or more exemplifications of
categorical universals. The most perspicuous way of understanding physical panpsychism via
this Armstrongian ontological take on the t-conception, then, is as attributing consciousness to
these exemplifications of causal relations. The relevant panpsychist claim is not that there is
something it is like to be an electron, but rather that there is something it is like to be an-
electron-causing-others-to-repel, or some such. The point in defence of premiss 3 still goes
through on this picture. These causal relation exemplifications are the physically basic items,
and an arbitrary world where all basic items are conscious and have their consciousness
properties fixed will thereby have its distribution of physical properties fixed. The key
ontological point remains the distinction between disposition exemplifications and their
categorical bases, even for theorists who deny that there is anything over and above an
electron’s categorical properties that determine its causal profile at a given world.

To the extent that the present defence of premiss 3 is cogent, argument 1-6 provides
new reason to doubt the conceivability of zombies.

3. Saving the phenomena: a test case

Why then do many theorists take zombies to be ideally positively conceivable? For the
purpose of examining a typical zombie thought experiment, let us stipulate temporarily that
panpsychism is false and that only complex macro objects are capable of phenomenal
consciousness. Suppose that two whitetail deer are such that their mereologically basic parts
are physically isomorphic. However, one of them, Bambi, is experiencing the taste of berries
and the other, Zambi, lacks any phenomenal experience.
Call the state of affairs described in the thought experiment ‘MicroDup’ since it fixes Bambi and Zambi as microphysical duplicates. Let us assume that the lower bound of Bambi’s MTC is located at the mereological level of some modest proper part of her brain, which serves as one of her mereologically minimal consciousness-capable proper parts. Call this brain-part ‘Min’ and the microphysically duplicate part of Zambi ‘Z-Min’. Consider now a state of affairs exactly like MicroDup except that in it Zambi is missing an important collection of quarks. The absence of this collection precludes the existence of Z-min or any other duplicates of any of Bambi’s consciousness-capable parts, but yields a consciousness-incapable near-duplicate of Min that we may call ‘Z-min*’, and thus a consciousness-incapable near-duplicate of Bambi. Call this second state of affairs that is exactly like MicroDup except that in it Zambi has no duplicates of any of Bambi’s consciousness-capable parts ‘MicroDupe’.

The problem for advocates of the original Bambi/Zambi example and its ilk is that it is not clear whether the temptation to endorse it as conceivable is due solely to the physical similarity between MicroDup and MicroDupe, rather than being linked to any legitimate coherence of supervenience-failure. This is a problem because Zambi is not a zombie in MicroDupe. She is merely a duplicate of one of Bambi’s mereologically maximal (or near maximal) consciousness-incapable undetached proper parts. Those who conclude by introspection that zombies are conceivable do so because they mistake states of affairs relevantly like MicroDupe for states of affairs relevantly like MicroDup.

Of course, an ideal conceiver would never mistake MicroDupe for MicroDup. But this does not show that zombies are conceivable. As we saw in section 1, an ideal conceiver would have such broad and sensitive knowledge of the correlation between Bambi’s physical parts
and the actual capability of being conscious (just as we non-ideal conceivers have coarse-grained knowledge of this correlation) that she might well reject as confused the suggestion that Zambi is unconscious in MicroDup. To the extent that we are tempted to think otherwise, we betray our ignorance of the fine-grained differences between actually consciousness-capable physical structures and their mereologically maximal consciousness-incapable undetached proper parts.

An important clarification is in order. The present diagnosis of the zombie theorist’s error does not require that the boundaries of MTC’s be single points on a given object’s mereological spectrum. It may be instead that there is some extended sub-threshold of borderline cases between parts that clearly are consciousness-capable and those that clearly are not. The diagnosis still goes through. All that is needed is a difficult-to-recognize cut-off between the clear cases and the borderline cases. Even if one were to deny that it is plausible that a single quark or relatively small collection of quarks could make a difference with respect to consciousness at the actual world, one need not deny that such fine-grained differences mark borderline from non-borderline cases. And even if one were concerned with hierarchies of meta-borderline cases, at some point in the hierarchy the mereological differences between distinct states of affairs would be sufficiently fine-grained as to obscure, for the non-ideal conceiver, which state of affairs plays which role in a given thought experiment.
4. Objections and replies

4.1 T-conception objection redux

In defence of premiss 3 of argument 1-6, you suggest disposition exemplifications as the most perspicuous bearers of consciousness for the t-conceptual treatment of conscious basic physical items; and you are neutral as to whether these disposition exemplifications are tropes or universals-at-locations (including second order universals-at-locations to accommodate Armstrongian ontologies of dispositions). But neither tropes nor universals are eligible to be conscious. Ascribing consciousness to them is a category mistake.

Response: The claim made in defence of premiss 3 is conditional in a way that this objection understates. The claim is that if the most basic physical elements in a given spatiotemporal-cum-mereological structure are conscious then the most perspicuous t-conceptual treatment is to ascribe consciousness to the most mereologically basic disposition exemplifications. Friends of the t-conception who reject this approach and instead ascribe consciousness to the objects that exemplify the categorical properties that ground the dispositions fail to ascribe consciousness to physical items and thereby fail to respect the antecedent of the conditional claim. Ascribing consciousness to the disposition tropes or located universals themselves is thus the most promising strategy by default. Those who charge this strategy with making a category mistake simply motivate the o-conceptual treatment of the conditional’s antecedent. Moreover, given how unfamiliar mereologically basic conscious items are to ordinary experience, it is far from clear that they must be traditional substances instead of tropes or universals-at-locations. The charge of category mistake is difficult to validate in conceivability debates involving consciousness.
4.2 The *tu quoque* objection

If the bounds of an arbitrary actually conscious object’s MTC can conceivably move anywhere along the object’s mereological spectrum, then it is conceivable that the upper bound (i.e. the point on the spectrum that contains the mereologically maximal consciousness-incapable proper part(s)) contains the object itself: a zombie. So the reasoning that leads to premiss 3 of argument 1-6 also leads to the conceivability of zombies.

Response: The first part of the argument for premiss 3 of argument 1-6 is that there is as good a conceivability argument for conscious mereologically basic physical items as there is for zombies. This is a problem for friends of zombies because conceivability arguments are the primary basis for zombie endorsement. This is not the case for physical panpsychism, which may be independently defended for its theoretical power, elegance, unity, etc. So, while it is a problem for friends of zombies that their conceivability standards lead equally to physical panpsychism and zombies, it is not a problem for physical panpsychists. Moreover, physicalists in general are free to deny that MTC’s are contingent, in which case the *Tu Quoque* Objection does not apply. Physicalists will be happy to end argument 1-6 by taking the latter disjunct of premiss 2.1, precluding the need for premiss 3.

4.3 The *NFM objection*

Some theorists (Wilson 2006) will deny the (ideal positive) conceivability of physical panpsychism because they hold, through independent argument, a ‘no fundamental mentality’ thesis (NFM) according to which any adequate understanding of the physical will preclude
fundamental mental properties. If NFM is true then physical panpsychism is incoherent and the consequent of premiss 3 of argument 1-6 is necessarily false.

Response: Either the NFM theorist will be a physicalist or not. If so, she will be happy to end argument 1-6 by taking the latter disjunct of premiss 2.1, leaving her dissatisfaction with premiss 3 inconsequential. If not, she will allow the conceivability of an idealist panpsychism that is parallel to physical panpsychism in all presently relevant respects. According to this idealist version, the fundamental properties are strictly mental even though they underwrite all properties required by current or ‘final’ physics. Idealist panpsychism works just as well in argument 1-6 as does physical panpsychism. The only difference is whether the fundamental consciousness properties count as compatible with physicalism or not, but this makes no difference with respect to being inconsistent with the possibility of zombies. 14 So whether or not its proponents are physicalists, NFM has no bearing on the soundness of 1-6.

4.4 The non-zombies objection
Conceivability arguments against physicalism (Chalmers 2002a, 2010a) do not turn on the conceivability of zombies per se, but merely on the conceivability that the physical facts do not strictly determine the phenomenal facts. Zombies are one putative counterexample, but partial zombies (physical twins who only partially lack consciousness) and inverts (physical twins with varied specific phenomenal states) work just as well against physicalism. So even if zombies are inconceivable, conceivability arguments from partial zombies and inverts still go through against physicalism.

14 Chalmers (2002a) makes this point.
Response: Argument 1-6 works just as well against partial zombies as it does against full zombies. A partial zombie has all the same physical proper parts as its fully conscious counterpart, but *ex hypothesi* it does not have all the same consciousness-capable parts. So it is inconsistent with both the necessity of MTC location and physical panpsychism. The case is not as straightforward for inverts, but a plausible tweak to argument 1-6 allows it to work against them as well. Just as an ideal conceiver’s physical knowledge enables her to fill in arbitrary details about consciousness-capability for fine-grained actual physical structures (which are assumed not to be zombies), so it enables her to fill in arbitrary details about consciousness specificity for fine-grained actual physical structures (which are assumed not to be inverts relative to each other). For example, she would be able to deduce exactly which structures are sufficient for outward signs (whether functional, behavioural, or causal) of specific phenomenal states, giving her extremely fine-grained information linking structure to specific phenomenal states. So the ideal agent’s knowledge of MTC location for arbitrary actually conscious object x will be augmented with information about the specifics of the conscious states of which x’s parts are capable. But once this is granted, argument 1-6 goes through just as well when ‘zombies’ in premisses 1 and 6 is replaced with ‘inverts’. Physical panpsychism is inconsistent with inverts, and the latter disjunct of premiss 2.1 renders inverts incoherent, given that facts about MTC’s now include phenomenal specifics. After all, if MTC location and phenomenal specifics are necessary (per the latter disjunct of 2.1), then no microphysical duplicate of an actual object in phenomenal state Q could fail to be in Q.
4.5 The ‘I just can’ objection

Look, I understand the difference between MicroDup and MicroDupe perfectly well, and I know that I can conceive of MicroDup-plus-unconscious-Zambi just by trying.

Response: This objection misses the point of ideal inconceivability. Consider again the sofa example from section 1. Smith owns a certain sofa that is barely too big to fit through the doorway of a prospective apartment. Call the state of affairs described in the example ‘SofaBig’, since the sofa is prohibitively big. Consider a second state of affairs, ‘SofaSmall’, which is indiscernible from SofaBig except that in it Smith’s sofa is just small enough to fit through the doorway. Assume that Smith is ignorant of the precise widths of the sofas and doorway in each state of affairs. He is ignorant of the widths, but perfectly able to understand what it is for them to be different widths. If we were to teach Smith all the present terminology without telling him which state of affairs is actual, he might well mimic the proponent of the ‘I just can’ objection: ‘Look, I understand perfectly well what the difference is between SofaBig and SofaSmall, namely, a few millimetres in respective sofa length. Still, I can conceive of my sofa [where his utterance of ‘my sofa’ is an indexical reference to the sofa exactly as it is in SofaBig] fitting through that doorway [demonstrating the doorway in SofaBig].’ But Smith is confused. SofaBig-plus-fitting contradicts actual physics. The worry for those who endorse the conceivability of zombies is that they are mistaken in a way analogous to Smith.
4.6 Chalmers's type-C objection

According to David Chalmers (2012), ideal inconceivability presupposes what he calls ‘type-C materialism’, the view that an ideal understanding of the nature of consciousness would close the epistemic gap between the physical and the phenomenal, revealing that zombies are inconceivable. But Chalmers (2002a) has already argued as follows that type-C materialism is untenable:

1. Physical truths concern only structure and dynamics.
2. Only further truths about structure and dynamics are deducible from truths about structure and dynamics.
3. Truths about consciousness are not truths about structure and dynamics.

So, even for an ideal cognitive agent, truths about consciousness cannot be deduced from physical truths.

Response: Chalmers’s (2002a) argument for his premiss 3 is uncompelling. He defends premiss 3 in three ways. First, he cites the knowledge argument (Jackson 1982). Mary is omniscient with respect to physical facts about colour but ignorant of what it is like to experience colour phenomenally; so phenomenal experience outstrips the physical. But it would be a significant concession if the conceivability of zombies turned on the independently controversial knowledge argument. Moreover, the physicalist who endorses ideal inconceivability is unlikely to be moved by the knowledge argument. Just as an ideal conceiver would catch a coherence-blocker that we do not when we think we conceive of zombies, so she would catch a coherence-blocker that we do not when we think we conceive of Mary. For example, such a physicalist will not want to rule out the suggestion that when
one ideally grasps the fundamental microstructure of mental states that feature red qualia (as Mary would be able to do), one actually has a corresponding internal experience with a red quale. This is far-fetched, perhaps, but it is just one case that the ideal inconceivability-endorsing physicalist is free to use to flesh out the contention that an ideal conception of the Mary thought experiment would betray incoherence.

Chalmers’ second claim in defence of premiss 3 is that for any structural or dynamical state, one can conceive of its instantiation sans consciousness. But this is unhelpful since it is the very case that the advocate of ideal inconceivability rejects.

Chalmers’ third and final claim in defence of premiss 3 is that

…to resist [premiss 3], an opponent would have to hold that explaining structure and dynamics thereby suffices to explain consciousness. The only remotely plausible way to do this would be to embrace type-A materialism [according to which there is in fact no deep epistemic gap between physics and consciousness]…’ (Chalmers 2002a, p. 259, italics in original).

Chalmers (2002a) argues at length against type-A materialism.

But Chalmers’s description of his opponent’s commitment here is ambiguous. What does it mean to ‘explain consciousness’? If it means to incite a conceptual revolution that illuminates some unknown features of the nature of phenomenal consciousness and its relation to the physical world, then the quoted claim is plausible. And it may be that Chalmers’s type-C materialist is committed to such an epistemically robust explanatory link between structure and consciousness. But if so then the present essay is not best construed as
advocating type-C materialism, and Chalmers is too quick in classifying ideal inconceivability as a marker of type-C materialism (or perhaps the taxonomy needs some sub-distinctions). If, by contrast, the claim to ‘explain consciousness’ via structure and dynamics is construed as the claim that ideal knowledge of physical structure will entail knowledge of whether consciousness is instantiated by an arbitrary structure, then the sort of knowledge of structure ascribed to the ideal conceiver in this essay is capable of ‘explaining consciousness’ independently of type-A materialism. The ideal conceiver described here will find zombies inconceivable even if she does not fully bridge the epistemic gap.

This is because the sense of explanation for the ideal conceiver described here is ontological, not epistemic. It does not purport to close the epistemic gap (though the question of whether it might in fact help to close it remains live). Rather, it purports to articulate the complete catalogue of exactly those actual physical structures that are sufficient for consciousness-capability across all physically resembling states of affairs. We may understand this catalogue as containing two lists. The first lists all actually consciousness-capable physical structures. The second lists all non-actual states of affairs in which the structures on the first list are consciousness-capable. If argument 1-6 of section 2 is sound then this second list will be attainable from the first and will contain all states of affairs that are physical duplicates of the actual world. With possession of this catalogue, one can deduce that zombies are incoherent even if one leaves open certain questions about why the catalogue contains exactly the items it does.

A final analogy may be helpful. Some United States pennies bear a letter near the date marker on the obverse side, denoting the location at which they were minted. Denver (‘D’) is the most commonly denoted, though there are also mints in Philadelphia (‘P’) and a few other
locations, none of which have a name that begins with ‘Z’. Consequently, there are no U.S.
pennies (as of early 2012) that bear a ‘Z’ on the obverse side. At any rate, let us assume as
much for the sake of this example. Consider now a state of affairs S that is physically
indiscernible from the actual world at the present moment. Given that no U. S. pennies bear a
‘Z’ on the obverse side at the present moment, it is inconceivable for an ideal agent that S
contains a U. S. penny that bears a ‘Z’ on the obverse side. For short, let us say that ‘S-plus-Z-
pennies’ is inconceivable. Now suppose that I know nothing about the system for denoting
U.S. mint locations on pennies; I simply know that some U.S. pennies have a single letter near
the date marker on the obverse side. My ignorance affords something of an epistemic gap for
me between an arbitrary coin’s being a U.S. penny and its bearing a certain letter near the date
marker on the obverse side. Compare a conceiver who has a complete and accurate catalogue
of extant U.S. pennies, including facts about which pennies bear which letters, but who is as
ignorant as I am supposing myself to be of the system for denoting mint location on pennies. I
may well find S-plus-Z-pennies conceivable even though the conceiver in question would not.
The conceiver with the complete penny catalogue would have an ontological explanation for
the inconceivability of S-plus-Z-pennies. But, and here is the point of the analogy, the
conceiver in question would not be able to close the epistemic gap between an arbitrary coin’s
being a U.S. penny and its bearing a specific letter on the obverse side, for in order to close
that gap she would have to know, inter alia, that the letters abbreviate names of U.S. mint
locations and that no U.S. mint locations have a name beginning with ‘Z’. Knowledge of
these facts is not entailed by knowledge of the catalogue that provides the ontological
explanation for the inconceivability of S-plus-Z-pennies.
Similarly, knowledge sufficient to close the epistemic gap between consciousness and physical structure may not follow from one’s possession of a complete catalogue of the myriad actual physical structures that are sufficient for consciousness-capability across all physically resembling worlds, for such a catalogue need not contain explanatory criteria for its precise contents. Nevertheless, any agent in possession of this catalogue will reject zombies as inconceivable since all actually conscious physical structures will be listed in it, entailing that they have no possible unconscious counterparts in physically duplicate worlds. If the arguments above are sound then the ideal agent knows the catalogue well.\footnote{Thank you to Alex Skiles, Elanor Taylor, Donovan Wishon, and the editor of this journal for helpful comments on ancestors of this essay.}
References


