

WHY ZOMBIES ARE INCONCEIVABLE

Eric Marcus

I argue that zombies are inconceivable. More precisely, I argue that the conceivability-intuition that is used to demonstrate their possibility has been misconstrued. Thought experiments alleged to feature zombies founder on the fact that, on the one hand, they *must* involve first-person imagining, and yet, on the other hand, *cannot*. Philosophers who take themselves to have imagined zombies have unwittingly conflated imagining a creature who lacks consciousness with imagining a creature without also imagining the consciousness it may or may not possess.

Thought-experiments have played a crucial role in the recent resurgence of interest in consciousness. Looming large in recent philosophical discussions of consciousness are claims that we can imagine creatures physically and functionally identical to us, but who have no consciousness—‘zombies’. Physicalists and functionalists are the most direct targets of such thought experiments, which threaten to establish the possibility of what are, according to these views of the mind, impossible. Without committing myself to either of these zombie-sensitive views, I will here call into question arguments that purport to establish the possibility of zombies. My claim will not be that zombies are impossible, just that the conceivability-intuition that is used to argue for their possibility has been misconstrued.

There are two ways of resisting arguments from conceivability or imaginability¹ to metaphysical possibility (hereafter, simply ‘possibility’). One can deny that what is alleged to be conceivable is in fact conceivable [Botterell 2001; Dennett 1995; Kirk 1999; and Shoemaker 1999]; or one can deny that conceivability establishes possibility [Botterell 2001; Balog 1999; Brueckner 2001; Hill 1997; Levine 2001; and Yablo 1999]. In what follows, I pursue the former strategy. Zombie worlds, I will argue, are not imaginable. In brief, I will argue this: Although there is a sense in which we can imagine a physical and functional duplicate of our world without consciousness, ‘without consciousness’ modifies the *act*, rather than the *object* of imagining. It’s not that we imagine a physical and functional duplicate of our world that contains no consciousness, but rather we imagine a world physically and functionally

¹ In what follows I will use the words ‘imaginable’ and ‘conceivable’ (and their cognates) interchangeably.

identical to ours without imagining the consciousness that it may or may not contain. The conceivability-intuition at issue is thus based not on our ability to positively imagine a zombie world, but rather on our ability to imagine a world physically and functionally like ours while refraining from imagining the consciousness it may or may not contain.

In order to better position my argument with respect to recent work on conceivability, I will begin, in section I, by exploring the taxonomy of kinds of conceivability recently canvassed by David Chalmers. In section II, I argue, following the strategy just outlined, that zombie worlds are not conceivable—notwithstanding Chalmers’s claim that ‘almost everybody’ can conceive of them. Chalmers holds that zombie worlds have what he calls ideal positive primary conceivability, and that this sort of conceivability provides the strongest support for possibility-claims. In section III, I will, using the argument of the previous section, show that zombies lack ideal positive primary conceivability. It is, of course, the adherents of zombie-sensitive views that have the most to gain from my argument. But, as I will ultimately suggest, regardless whether one holds such a view or not, the inconceivability of zombies might also allow substantial headway on the infamous problem of the explanatory gap between the objective and subjective aspects of the world.

I

Chalmers distinguishes between *prima facie* and ideal conceivability, positive and negative conceivability, and primary and secondary conceivability. Zombies, he claims, have ideal positive primary conceivability. Let us briefly go over each part of this characterization.

A statement *S* is ideally conceivable if it is conceivable on ideal rational reflection. It is *prima facie* conceivable for a subject if it is ‘conceivable for that subject on first appearances’ [Chalmers 2002: 147]. A hypothesis might be *prima facie* conceivable for a subject and yet ideally inconceivable if some failure of attention, intelligence, or clarity of thought creates the false appearance of possibility. I will ultimately accuse the advocate of the possibility of zombies of a failure to think clearly through the alleged possibility. Thus, I will argue that it has at best *prima facie* conceivability—only the mistaken appearance of conceivability.

S is negatively conceivable² ‘when *S* is not ruled out a priori, or when there is no (apparent) contradiction in *S*’ [ibid.: 149]. The negative conceivability of

² This is what Chalmers calls ‘the central sort’ of negative conceivability; the non-central sorts need not concern us here.

a claim is entailed by, but does not entail, its positive conceivability. Positive conceivability requires not just the absence of contradiction, but also ‘that one can form some sort of positive conception of a situation in which S is the case’ [ibid.: 150]. Distinctive of positive conceivability is what Chalmers calls its *mediated objectual character*. When I imagine a purple spittoon, I might form a mental image that is *of* a purple spittoon. The image is an intermediate mental object between the possibility and me.

This is an instance of perceptual positive imagining—the mediating object is a visual image. But, according to Chalmers, positive imagining need not be perceptual. I can imagine what is ‘beyond the scale of perception: for example, molecules of H₂O or Germany winning the Second World War’ [ibid.: 151]. In such cases, there are no mental images, instead ‘we have an intuition of (or as of) a *world* in which S, or at least of (or as of) a situation in which S, where a situation is (roughly) a configuration of objects and properties within a world’ [ibid.]. In the case of *modal* imagining, then, the mediating object is an *intuition* that represents the possibility in a way analogous to that of the image in the case of perceptual imagining.³

It will be worthwhile for the discussion in the following section to introduce a different sort of example of modal imagining, a sort Chalmers does not consider in this connection. I can positively imagine someone feeling exhilarated upon watching a golf tournament, though I feel no such thing. I do, however, know what it’s like to feel exhilarated upon watching other sporting events. In this case, I imagine someone feeling towards golf the way I feel about those other sporting events. Here we have a case of what I call *first-person* imagining, as opposed to the *third-person* imaginings above (works that employ this or related distinctions include Nagel [1974] and Shoemaker [1994]). First-person imagining involves imagining what it’s like to have various experiences. In such cases, we have an intuition of a world or situation in which there is a certain configuration of objects and properties; but these objects and properties are understood from a subjective point of view.

Finally, we must distinguish between primary and secondary conceivability. Here’s how Chalmers does it: ‘We can say that S is *primarily conceivable* (or epistemically conceivable) when it is conceivable that S is *actually* the case. We can say that S is *secondarily conceivable* (or *subjunctively conceivable*) when S conceivably might have been the case’ [ibid.: 157]. Secondary conceivability is the more familiar sort of conceivability. It is secondarily conceivable that Harvard gives hacky-sack scholarships, but not that water contains no hydrogen. Any possible world that contains no

³ I don’t see exactly how an intuition could be an object analogous to an image; but I won’t digress by disputing that here.

hydrogen contains no water, even if it does contain a substance with the superficial properties of water, for water is identical to hydrogen dioxide. It is primarily conceivable that water contains no hydrogen, however, because, Chalmers claims, it conceivably might turn out that water doesn't contain hydrogen. Imagine—the chemists call a press conference, announce that there's been a terrible mistake, water only appeared to be H₂O because The primary conceivability of water containing no hydrogen consists in the conceivability of such scenarios.

I will not attempt to make the primary/secondary conceivability distinction precise.⁴ We are already in a position to see how this distinction would be of particular relevance for arguments from conceivability to possibility. Working with what Chalmers later dubbed secondary conceivability, Kripke showed that not everything we take to be conceivable is in fact conceivable. It might have *seemed* conceivable that water was a simple substance; but there is no possible world in which water contains no simpler parts. What we were in fact conceiving of was a world with a simple substance that has the superficial properties of water, but with no water. In the case of secondary or Kripkean conceivability, the link between conceivability and possibility remains tight. If you can conceive it, it's possible. But he weakened the link between apparent conceivability and conceivability. Whether or not certain hypotheses are conceivable—secondarily conceivable, in our newly refined language—depends upon facts about which we may have no knowledge, facts about the constitution of natural kinds, for example.

Part of Chalmers's interest in primary conceivability derives from the fact that knowledge of primary conceivability can be entirely a priori; that is, whether or not a certain hypothesis is primarily conceivable is a matter that could be revealed by rational reflection alone. I know a priori that six could not turn out to be nine; and I know a priori, Chalmers would argue, that water could turn out to be XYZ. (In general, the falsity of any of Kripke's necessary a posteriori identities is secondarily inconceivable, but primarily conceivable.) Thus there is a strong link between apparent primary conceivability and primary conceivability.

We should next consider the link between primary conceivability and possibility. Chalmers does not claim that the primary conceivability of 'water is XYZ' corresponds to a world in which water is XYZ, for there is no such world. 'Water is XYZ' is possible—primarily possible, that is—but its primary possibility consists in the fact that there is a possible world where something that isn't water, but which looks and acts like water, is called 'water'. Primary conceivability is thus strongly linked to primary possibility;

⁴ I don't see exactly how to make this distinction precise; but it is not my burden here to do so.

and primary possibility involves the possibility of a world; but the world may not be properly described by the language of the relevant hypothesis. 'Water is XYZ' is primarily conceivable; and this guarantees the existence of a possible world, just not a world in which water is XYZ.⁵

What, then, of the ideal positive primary conceivability of zombies? If primary conceivability does not guarantee the existence of a possible world properly described in the language of the hypothesis, then why should the primary conceivability of zombies lend support to the idea that there *is* a possible world in which they exist? Chalmers's Kripkean answer to this question is as follows: A world that lacks water but has XYZ can be mistaken for a watery world only because XYZ presents itself to us in just the way H₂O does—they have the same superficial properties. But a world in which nothing presents itself to anyone as, for example, pain—a world in which nothing possesses the superficial features of pain—is a world without pain. More generally, a world that lacks the superficial features of consciousness is a world in which there is no consciousness. If zombies are primarily conceivable, then there is a world that is either itself a zombie world or is a superficial twin of a zombie world. But a superficial twin of a zombie world is also a zombie world. That zombies are primarily conceivable thus shows there is a world in which zombies exist.

It is worth pointing out that Chalmers has not attempted to *define* conceivability *simpliciter*. Chalmers had defined three dimensions along which kinds of conceivability vary. But the definitions of *prime facie* vs. *ideal*, and *primary* vs. *secondary* conceivability contain the word 'conceivable' on the right-hand side. In these definitions at least, the core notion of conceivability remains an intuitive one. In what follows, when I use the word 'conceivable' without any of these just-discussed modifiers, I am using it in the intuitive sense in which it is used on the right-hand side. In the next section, I argue that the zombie hypothesis is not conceivable *simpliciter*. In section III, I will situate the conclusion of this argument in Chalmers's taxonomy of conceivability.

II

Why does Chalmers take zombies to be conceivable? Here's what he says

I confess that the logical possibility of zombies seems equally obvious to me [as that of a mile-high unicycle]. A zombie is just something physically

⁵ Though no substantive point turns on this way of putting it, Chalmers would prefer to put it differently. The world that corresponds to the primary conceivability of 'water is XYZ' is not described by the language of the hypothesis only if this language is understood in terms of secondary, rather than primary, intensions.

identical to me, but which has no conscious experience—all is dark inside. While this is probably empirically impossible, it certainly seems that a coherent situation is described; I can discern no contradiction in the description. In some ways an assertion of this logical possibility comes down to a brute intuition, but no more so than with the unicycle. Almost everybody, it seems to me, is capable of conceiving of this possibility.

[Chalmers 1996: 96]

In this section, I will attempt to dislodge this intuition by (a) arguing that zombies are inconceivable and (b) diagnosing the contrary intuition.

Let us go over carefully what we are supposed to imagine.

Zombies are physically, functionally, and behaviourally identical to us.⁶ More generally, zombies are supposed to have every property of ours that can, in some important sense, be reductively explained in terms of basic physics. And it is typically thought that the only properties that may resist this sort of explanation in the end are the essentially subjective properties, properties that are graspable only from a first-person point of view. In fact, the conceivability of zombies is supposed by many to establish that subjective properties, unlike other properties, cannot be reductively explained in physical terms. (They are thought to be the only properties in relation to which there arises an ‘explanatory gap’.) It is essential to the point of these thought experiments, then, that there is nothing, from the point of view of an observer of a zombie, that can distinguish it from a person. Let us put this by saying that we are to imagine beings that are *third-personally* or *objectively* identical to ordinary, conscious beings, yet without consciousness.

The imagined difference between zombies and us is purely first-personal or subjective. What is such a difference? As typically understood, a first-person difference is a difference in what it’s like To imagine Abe Lincoln third-personally is to imagine, for example, the way his parts are laid out in space. To imagine him first-personally is to consider how the world appeared to him, to imagine feeling what he felt, experiencing his moods, and so forth. To imagine a subjective difference between Abe Lincoln and someone else is to imagine a difference in what it’s like to be them.

To imagine zombies first-personally, then, is to imagine what it’s like to be a zombie. What is it like to be a zombie? Well, zombies are beings without consciousness. There is *nothing* that it’s like to be a zombie. To imagine zombies first-personally is to imagine first-person nothingness. The difference between Abe and Zombie-Abe is that there is something it’s like to be

⁶ Some have argued that the idea of zombies is incoherent in virtue of the clash that would develop between conscious states and other mental states. See, for example, Balog [1999]. These arguments, however, depend on the view that zombies would still have other mental states. I do not make this assumption here.

Abe, and nothing it's like to be Zombie-Abe. So, when we add third-person duplication to first-person absence, we've imagined zombies.

But what is it to imagine first-person absence? Is third-person absence imaginable? It seems to be: empty rooms, empty heads, houses with no one home, and vast stretches of Nebraska are imaginable. In general, empty space is imaginable. In these cases, however, absence is imaginable against the background of presence—the presence of rooms, heads, houses, cows, and, in general, space. But there is nothing comparable in the case of the nothing it's like to be a zombie. There is no inner border or background of inner space against which it is possible to conceive subjective absence. Imagining subjective absence presents an insurmountable obstacle. On the one hand, it is something that we are to imagine from the first-person point of view—it's not an objective matter. It is supposed to be a subjective fact about zombies that they lack consciousness. On the other hand, there is nothing that it's like to be subjectively absent. So there is no imagining of the what-it's-like variety that we can use to arrive at this possibility.

I submit that the barrier to imagining zombies is much higher even than the barrier to imagining how a bat feels. The integration of echolocators into our sensory-motor apparatus might at least begin to make such imaginings possible. In the case of zombies, however, there is no superior capacity or new sensory faculty whose possession will enable us to imagine what it's like. The problem for this thought-experiment is not that there is something imaginable, only we can't quite conjure it up. The problem is that there is nothing to be imagined. To 'imagine' creatures that are objectively identical to us with all subjectivity removed is neither an act of third-person imagining, nor an act of first-person imagining. No, to 'imagine' a zombie is not really to imagine at all.

It is worth emphasizing that it is crucial for the argument in favour of the possibility of zombies that the conceivability of zombies be in part a matter of first-person imagining. As I hope is already clear, no purely third-person imagining by itself will conjure up the alleged zombie-world. That what needs to be added to third-person imagining is first-person imagining can be seen by considering the way Chalmers's argument is supposed to parallel Kripke's argument against the identity thesis. Kripke's assertion that any world that seems to have pain, has pain, only succeeds because pain is being imagined from the first-person point of view. His argument works because anything that *feels like pain* is pain. Similarly, Chalmers's assertion that any world that seems to have no consciousness is a world that in fact has no consciousness will succeed only if the absence of consciousness is imagined from the first-person point of view. His argument only appears to work because anything that *doesn't feel like pain* isn't pain. And the trouble for Chalmers is that while the absence of pain specifically is subjectively imaginable, the absence

of consciousness altogether is not. We saw above that the primary possibility of water being XYZ does not secure the existence of a world where water is XYZ. For the primary conceivability of zombies to secure the existence of a world that contains zombies, the primary conceivability of zombies must in part be a matter of first-person imagining. But it can't be.

If zombies are unimaginable, why do so many believe that they have imagined them? I suggest that those who take themselves to imagine zombies are mistaking not imagining something for imagining nothing. To imagine creatures that are third-personally identical to persons, and first-personally absent is less to imagine than to refrain from imagining, namely to refrain from imagining their inner lives. We imagine creatures just like us, but we diligently refrain from imagining how things are for them. But not to imagine and to imagine nothing are distinct. The alleged possibility of zombies, I contend, is not something we arrive at by imagining, but rather something we arrive at by refraining from imagining.

Now, our ability to imagine cuts some metaphysical ice. But our ability to refrain from imagining cuts no metaphysical ice. It does not follow from the fact that we can imagine creatures physically, functionally, and behaviourally identical to us without imagining their experiences that there can be creatures physically, functionally, and behaviourally identical to us who experience nothing. No one would argue that to imagine a happy family without imagining their toes is to imagine a toeless happy family. Similarly, it does not follow from the fact that we can imagine creatures third-personally like us without thereby imagining what it's like to be them, that we have imagined creatures third-personally like us whom there's nothing it's like to be. The possibility of zombies is supposed to follow from our ability to imagine them. Since no such ability has been demonstrated, no such possibility has been proven.

Consideration of a few objections will help to clarify the argument. A natural first response goes like this:

You say we cannot imagine these non-conscious beings; but we encounter and conceive of non-conscious beings all the time. Rocks, toasters, one-celled organisms, very small dogs, and unconscious people, just to name a few. In imagining the possibility of zombies, we simply impute to creatures that look and act exactly like us the property we commonly attribute to these other things. Why should that be so difficult?

It is of course true that one can conceive of unconscious beings. But we cannot simply infer the coherence of the mental act whereby one allegedly transfers to conscious people the property of unconsciousness from the fact that unconsciousness is conceivable. For the unconscious beings of which we can uncontroversially conceive are physically, functionally, and behaviourally

different from us. What makes it possible to conceive of such beings may well be essentially tied to conceiving also of one or more of these differences. We can only transfer unconsciousness to conscious beings, while otherwise leaving them intact, so to speak, if indeed the property of consciousness is not metaphysically necessitated by physical, functional, or behavioural facts. And that is just what the conceivability of zombies is supposed to settle. If my argument is correct, we thus cannot simply attribute to conscious beings the property uncontroversially attributed to unconscious beings.

The argument, together with the accompanying diagnosis, can be resisted more compellingly along the following lines:

You have placed a lot of weight on the idea that we have subtracted subjective properties. But we could just as well have framed the thought experiment as one whereby subjective properties are added. The form of the thought experiment is then as follows. Imagine a scenario *s*, where *s* corresponds to an objective description of the world. Now imagine a scenario *s* with additional feature *f*, where *f* corresponds to a subjective description of the world. The point of the thought-experiment is that one can imagine *s* without imagining *s* plus *f*. The difference between the two possibilities is easily graspable. In one case, the possibility is arrived at purely through third-person imagining; in the other, the possibility is arrived at through both first- and third-person imagining.

Does this approach save the thought-experiment? I don't think so. The details are important. It won't do to say: First, imagine a world that contains things that look and act exactly like people, yet lack consciousness; second, imagine a world that contains things that look and act exactly like people, and that also have consciousness. For the first step simply presupposes, on its face, the possibility of what I argued above was impossible. No, if the thought-experiment is really to have the form outlined by my critic, it will have to employ concepts that are already thoroughly objective, rather than concepts that require us to factor out the subjective. Instead of asking us to imagine creatures we ordinarily think of as conscious as lacking consciousness, a proponent of this method must begin with items we don't already think of as conscious.

A natural way of proceeding is as follows:

Step One: Imagine a world that contains molecules laid out exactly as our world's molecules are, without any consciousness.

Step Two: Imagine a world that contains molecules laid out exactly as our world's molecules are, with consciousness.

Here there is no problem of having to refrain from imagining the subjectivity of the molecules. There is instead, however, a problem of having to

imagine the subjectivity of molecules. Adding subjectivity to molecules—as we are asked to do in the Step Two of our new thought experiment—even in the imagination, is no simple matter. For we don't know how to imagine that molecules or sets of molecules are conscious—except insofar as we imagine molecules constituting complex living organisms (or facsimiles thereof). The problem is not just that there's nothing it's like to be a molecule, or a set of molecules, but that we have no idea what it would be like to be a molecule or set of molecules, even if it made any sense to suppose that they could be conscious.⁷

This consideration might seem irrelevant. After all, we still have imagined a physically identical universe with no consciousness—that's all that's required to get the possibility of zombies off the ground. Things are not that simple, however. It's important for the proponent of zombies that consciousness is something that's *added* in Step Two, in the sense of being a metaphysically independent extra. Indeed, that is just what's at issue. But consciousness is not obviously added in this sense.

Consider the following question:

In imagining Step One, are we to imagine (1) just the molecules, or (2) the molecules together with the things that the molecules compose?

Suppose we say (1). Which of the following is the best way to understand what we are to do:

(1a) We are to imagine a world that contains all of the molecules of a world, laid out exactly as our world's molecules are, but none of the mountains, planets, plants, etc.

(1b) We are to imagine a world that contains all the molecules, mountains, planets, plants, etc., but we are to refrain from imagining its mountains, planets, plants, etc . . .

If, as seems highly plausible, it is impossible for there to be a molecule for molecule duplicate of our world that lacks mountains, et al., then (1a) is not doable. (1b), however, is doable. But if (1b) is how we are to understand (1), then, whether we choose (1) or (2), the world that we are to imagine is the same: A world that contains not just our molecules but everything our molecules compose. But people are among the things that the molecules

⁷ It might be objected here that we can imagine the subjectivity of a molecule. Be that as it may, this fact cannot seriously be offered as one that underlies the alleged intuitiveness of thought experiments involving zombies. Cf., Wittgenstein [1953: §284]: 'Could one imagine a stone's having consciousness? And if anyone can do so—why should that not merely prove that such image-mongery is of no interest to us?'

in our world compose. People are not metaphysically independent extras in relation to their molecules—any more than mountains and planets are. To imagine a molecularly identical universe without imagining people is just to refrain from fully imagining a universe; it is to keep one's imaginative attention on the very tiny.

It will now be objected that I have begged the question. What is at issue just is whether people are metaphysically independent extras in relation to their molecules or not. That charge is misplaced however. Recall where we are in the dialectic. Above, I argued that we cannot imagine zombies by imagining people and then imagining that they lack consciousness. I offered a way of getting around this argument by starting with a molecule for molecule duplicate of our world and seeing whether we might be able to imagine it without consciousness. But if the world that we then imagine is one that contains everything that molecules compose, then we are back where we started. For molecules compose people. We thus return to the question of whether we can imagine creatures that are objectively identical to people, yet lack consciousness. And that question has already been answered.

I hold, then, that to the extent that the expression 'to imagine a physically identical world without consciousness' is genuinely intelligible, 'without consciousness' qualifies not an object of imagining, but an act of imagining. Not: We imagine p without q. But rather: We imagine p without imagining q.⁸

In the next section, I will consider how to fit this argument into Chalmers's taxonomy of conceivability set forth in the previous section.

III

If zombie worlds are not conceivable *simpliciter*, then, though they may be prime facie conceivable, they are not ideally conceivable. They are not ideally

⁸ An interesting objection that I do not consider, as it would take us too far from our topic here, is this. Many philosophers have worried that there is a problem of other minds. I know that I have an inner life. How can I know that others have an inner life as well? If zombies are inconceivable, however, then there really is no such problem, for there is no conceivable state-of-affairs for an argument against this kind of scepticism about other minds to rule out. But there is such a problem, my objector complains. I would prefer to see this consequence of the argument as a virtue rather than a vice. Virtue or vice, however, the conclusion of this paper doesn't put all worries about other minds to rest. It is, from my perspective, conceivable that that everyone I know lacks consciousness. And that is because all of the evidence is not in. My friends might all be a radio-controlled marionettes, for all I know. But this sort of local problem of other minds, as we might call it, will not restore the general metaphysical problem of other minds. That problem is: even if everyone looks to any observer, no matter how well-placed, just like a conscious being, perhaps there's really 'nobody home'. I have shown that there is no conceivable scenario that is raised by this worry.

conceivable because an inconceivable hypothesis would not be judged conceivable on ideal rational reflection. Since they are not conceivable *simpliciter*, it is not conceivable that one of them is actual. Hence, zombie worlds are not primarily conceivable. Nor is it conceivable that one might have been the case. Hence they are not secondarily conceivable.

But I have ignored the third dimension of difference between kinds of conceivability: What of positive and negative conceivability? I have shown at best only that zombie worlds are *positively* inconceivable, that we cannot coherently imagine a situation that, in Chalmers words, verifies the hypothesis that there are zombies (whether or not that hypothesis is considered as actual or counterfactual). To be *negatively* conceivable requires much less—just that S is not ruled out a priori, or that there is no (apparent) contradiction in S. And I have not argued that zombie worlds are ruled out a priori; nor have I derived a contradiction from the hypothesis that there are zombies. Though positive conceivability might make a stronger case for possibility than negative conceivability; the latter still provides some evidence, does it not? Perhaps.

Let us take negative secondary and negative primary conceivability in turn. ‘One might say that s is negatively secondarily conceivable when a priori reflection and empirical non-modal knowledge reveal no incoherence in the hypothesis that might have been the case’ [Chalmers 2002: 159]. But, as Chalmers is aware, ‘water isn’t H₂O’ would have satisfied this definition prior to the discovery that water is H₂O. This hypothesis was thus once negatively secondarily conceivable. It was never, however, secondarily possible, for there is no possible world where water is not H₂O. Nor was it ever even secondarily conceivable, for it is not conceivable that water might not have been H₂O. (So, a statement can be negatively secondarily conceivable, yet secondarily inconceivable!) In general, then, unknown necessary a posteriori falsehoods are all negatively secondarily conceivable impossibilities. Negative secondary conceivability is thus not a particularly good guide to possibility.

But Chalmers’s focus is on primary conceivability. ‘We can say that S is primarily negatively conceivable when it is not ruled out a priori that S is actually the case, or more briefly, if S is not ruled out a priori’ [ibid.: 158]. Is it ruled out a priori that the actual world contains beings functionally and physically identical to us, but who lack consciousness? The answer to this question is not obvious. In general, the best way to show that a statement is negatively conceivable is to positively conceive of it. That is, if I can coherently imagine a situation in which a certain state of affairs obtains, then I know that it is not ruled out a priori. But this is just what I can’t do in the case of the zombie-hypothesis.

Chalmers holds that negative primary conceivability is a good guide to (primary) possibility in large part because he thinks negative primary conceivability implies positive primary conceivability. But the zombie-hypothesis is not positively conceivable. So either Chalmers is wrong that negative conceivability implies positive conceivability (there are what he calls 'open conceivabilities' [ibid.: 186–8]) or he is wrong that zombies are negatively conceivable. Either way the argument for the conclusion that zombies are possible from their negative conceivability alone is weakened. If there are open conceivabilities, then the best reason for thinking that negative conceivability proves possibility is undermined. If zombies are not even negatively conceivable, then the last reason for thinking that zombies are possible has been undermined.

My view on the negative conceivability of zombies is therefore this: It is not obvious that they are ideally negatively conceivable; nor is it obvious that their negative conceivability provides a good reason for thinking them possible. But my point here is not to refute the argument from negative conceivability. I thus do not claim to have definitively refuted every conceivability argument for the possibility of zombies, just the best one. The negative conceivability argument remains in contention, but does not loom as the formidable threat to zombie-sensitive views that positive conceivability arguments have.

Conclusion

The problem of the explanatory gap between consciousness and the objective world is, as many have noted, peculiar and distinctive. Its peculiarity consists in the fact that it involves a demand for an explanation, with no conception of how any explanation could satisfy this demand. This has led some to predict that a conceptual revolution will be required in order to explain how consciousness arises from material stuff [Nagel 1974], and others to declare the human mind unfit for such an explanation [McGinn 1989]. Another reaction is to reject the demand for explanation as in some way ill-conceived from the start. But this reaction will appear *ad hoc* for at least as long we take ourselves to imagine how things might have been objectively just as they are, yet subjectively different. With apparent zombie and inverted-spectrum possibilities arrayed before the mind's eye, it seems that we can ask: Why not one of these other possibilities? If it could be shown that zombie worlds and inverted-spectrum worlds were not genuinely conceivable, however, then the contrast that bestows the appearance of intelligibility on the demand for explanation would

disappear [cf., Putnam 2000]. And the road would be paved for a conscientious rejection of the demand. My aim here has been to complete the first half of that task.⁹

Auburn University

Received: February 2003

REFERENCES

- Balog, Katalin 1999. Conceivability, Possibility, and the Mind-Body Problem, *Philosophical Review* 108: 497–528.
- Botterell, Andrew 2001. Conceiving What Is Not There, *Journal of Consciousness Studies* 8: 21–42.
- Brueckner, Anthony 2001. Chalmers’s Conceivability Argument for Dualism, *Analysis* 61: 187–93.
- Chalmers, David 1996. *The Conscious Mind*, New York: Oxford University Press.
- Chalmers, David 2002. Does Conceivability Entail Possibility?, in *Conceivability and Possibility*, ed. Tamar Szabo Gendler and John Hawthorne, Oxford: Clarendon Press: 145–200.
- Dennett, Daniel 1995. The Unimagined Preposterousness of Zombies, *Journal of Consciousness Studies* 2: 322–6.
- Hill, Christopher 1997. Imaginability, Conceivability, Possibility, and the Mind-Body Problem, *Philosophical Studies* 87: 61–85.
- Kirk, Robert 1996. Why There Couldn’t Be Zombies, *Supplement to the Proceedings of the Aristotelian Society* 73: 1–16.
- Levine, Joseph 2001. *Purple Haze*, New York: Oxford University Press.
- McGinn, Colin 1989. Can We Solve the Mind-Body Problem, *Mind* 98: 349–366.
- Nagel, Thomas 1974. What Is It Like To Be a Bat?, *Philosophical Review* 83: 435–50.
- Putnam, Hilary 2000. To Think With Integrity, *Harvard Review of Philosophy* 8: 4–13.
- Shoemaker, Sydney 1994. The First-Person Perspective, *Proceedings and Addresses of the American Philosophical Association* 68: 7–22.
- Shoemaker, Sydney 1999. On David Chalmers’s ‘The Conscious Mind’, *Philosophy and Phenomenological Research* 59: 439–44.
- Yablo, Stephen 1999. Concepts and Consciousness, *Philosophy and Phenomenological Research* 59: 455–63.
- Wittgenstein, Ludwig 1953. *Philosophical Investigations*, New York: Macmillan.

⁹ I would like to thank audiences at Auburn University and the North Carolina State Philosophical Association for helpful discussion. I also thank William Davis, Kelly Jolley, Roderick Long, Ram Neta, and Michael Watkins for helpful comments and suggestions.