

Imagining Subjective Absence: Marcus on Zombies

Torin Alter, talter@ua.edu

[Forthcoming in *Disputatio*]

Abstract. The claim that zombies are conceivable is a premise of one of the most important anti-physicalist arguments. Eric Marcus (2004) challenges that premise in two novel ways. He observes that conceiving of zombies would require imagining total subjective absence. And this, he argues, we cannot do. However, his argument turns on the assumption that absence is imaginable only against a background of presence and, I argue, that assumption is dubious. Second, he proposes that the premise's intuitive plausibility derives from a scope confusion. However, I argue, on reflection that proposal is untenable.

Many philosophers accept the conceivability of zombies: creatures that lack consciousness but are physically and functionally identical to conscious human beings. Many also believe that the conceivability of zombies supports their metaphysical possibility. And most agree that if zombies are metaphysically possible, then physicalism is false. So, the claim that zombies are conceivable may have considerable significance.¹

Eric Marcus (2004) challenges this conceivability claim two novel ways. First, he observes that conceiving of zombies (in the sense relevant to the anti-physicalist conceivability argument) requires imagining total subjective absence. And this, he argues, we cannot do. Second, he suggests that the plausibility of the claim that zombies are conceivable derives from a scope confusion.

Marcus' overall strategy has the right form. His proposals illustrate the kinds of substantive moves that would be required to answer the challenge to physicalism based on the conceivability of zombies.² But he does not succeed in this endeavor. His

argument against the possibility of imagining complete subjective absence turns on the assumption that absence is imaginable only against a background of presence. But, as I will argue, that assumption is dubious. And, as I will also argue, his scope-fallacy diagnosis of the intuition that zombies are conceivable is untenable.³

1. The conceivability of zombies

Marcus focuses on Chalmers' (2002) claim that zombies are *ideally primarily positively* conceivable. Let me explain this terminology.⁴ A hypothesis *S* is *ideally* conceivable if it is conceivable on ideal rational reflection. This contrasts with *prima facie* conceivability, or conceivability on first appearances. Marcus thinks that, while zombies seem conceivable, on reflection they are not.

S is *primarily* (or *epistemically*) conceivable if it is conceivable that *S* is actually the case. *S* is *secondarily* (or *subjunctively*) conceivable if *S* might have been the case—counterfactually, given how things actually are. Let S_w be the hypothesis that water lacks hydrogen. If Kripke (1972) is right, then S_w is not secondarily conceivable: given that water is actually H_2O , there is no counterfactual situation with hydrogen-free water, and so no such situation is secondarily conceivable. But S_w is primarily conceivable: a priori it might turn out that the watery stuff (the stuff that runs out of faucets, fills lakes and oceans, etc.) in the actual world contains no hydrogen; imagine that chemists reveal a massive hoax. Marcus' arguments concern the primary conceivability of zombies: the claim that a priori it might turn out that the actual world is a zombie world (a minimal physical/functional duplicate of the actual world but without consciousness).⁵

S is *negatively* conceivable roughly when S is not ruled out a priori. Positive conceivability is a stronger notion. S's *positive* conceivability (or *imaginability*) requires also "that one can form some sort of positive conception in which S is the case" (Chalmers 2002: 150). Positive conceivability has what Chalmers calls a *mediated objectual character*. For example, to perceptually imagine a situation in which pigs fly, we do not merely entertain a hypothesis. Rather, "we form a mental image that represents a specific situation (one with a certain configuration of animals), and we take this to be a situation in which pigs fly" (Chalmers 2002: 150). The mental image plays a crucial role in mediating between the conceiver and the hypothesis. But positive conceivability need not be perceptual. For example, we can imagine "molecules of H₂O, or Germany winning the Second World War" (Chalmers 2002: 151). Here intuition, not imagery, provides the mediated objectual character. To imagine H₂O molecules, we might imagine a configuration of particles. To imagine Germany winning WWII, we might imagine a situation in which "certain German armies win certain battles and go on to overwhelm Allied forces within Europe" (Chalmers 2002: 151).

2. Can we imagine subjective absence?

2.1 The entailment principle

Marcus argues that, on reflection, zombies are not (primarily) imaginable. He writes,

...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. (Marcus 2004: 483)

Here Marcus seems to assume what I will call *the entailment principle*:

(EP) To imagine the instantiation of a negative property N, one must imagine the instantiation of some positive property P that entails N.⁶

For example, to imagine hairlessness, we must imagine, say, a shiny scalp; to imagine a house with no one home, we must imagine, say, a house with just furniture and air occupying the relevant spaces; and so on. The zombie hypothesis also involves a negative property: the absence of consciousness. Thus, on the entailment principle, we can imagine zombies only if we imagine some positive property that entails the complete absence of consciousness. This positive property must also be a priori compatible with the idea that the imagined world is a minimal physical/functional duplicate of the actual world. But what positive property could serve this purpose? None that we can imagine, according to Marcus. So, he concludes, we cannot imagine zombies.

But why accept the entailment principle? Marcus does not explain, even though it (or something in the vicinity) seems pivotal to his main argument. Yet the principle is implausible. We can imagine an ectoplasm-free world without imagining a positive property that excludes ectoplasm; an angel-free world without imagining a positive

property that excludes angels; a Godless world without imagining a positive property that excludes God; and so on. At least such things seem imaginable. And Marcus provides no basis for doubting that they are.

I have formulated the entailment principle in highly general terms. Perhaps Marcus relies on a more restricted version. He repeatedly (and correctly) emphasizes that conceiving of zombies requires *first-person* imagining. He writes,

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. ...no purely third-person imagining by itself will conjure up the alleged zombie-world. (Marcus 2004: 483)

So, perhaps he relies on a principle that pertains specifically to first-person imagining, such as:

(EP') To imagine, from the first-person perspective, the instantiation of a negative property N, one must imagine, from that perspective, the instantiation of some positive property P that entails N.

But EP' seems no more plausible than EP. Why can't we imagine, from the first-person perspective, a pain-free world without imagining, from that perspective, a positive property that excludes pain? An elation-free world without imagining, from that perspective, a positive property that excludes elation? A world without auditory

experiences without imagining, from that perspective, a positive property that excludes such experiences? If there is an insurmountable obstacle to imagining such things, it is hard to see.

Marcus contrasts imagining zombies with Kripke's case of imagining C-fiber stimulation without pain. Marcus writes, "while the absence of pain specifically is subjectively imaginable, the absence of consciousness altogether is not" (Marcus 2004: 483-84). This yields a yet more restricted version of the entailment principle:

(EP'') To imagine, from the first-person perspective, the instantiation of a negative property N, one must imagine, from that perspective, the instantiation of some positive property P that entails N. Also, if N concerns consciousness, then N must be an absence of a specific kind of sensation, such as pain, rather than the absence of consciousness in general.

But why does specificity matter? Marcus provides no good reason. And the counterexamples I adduced against EP' also seem to refute EP''. For example, it seems that we can imagine, from the first-person perspective, a specifically pain-free world without imagining, from that perspective, a positive property that excludes pain.

Marcus' contrast between imagining the absence of pain and imagining the total absence of consciousness suggests that he accepts the claim that we can imagine a minimal physical/functional duplicate of the actual world but without pain. But the latter claim strengthens the case for the imaginability of a zombie world. If we can imagine a minimal physical/functional duplicate of the actual world but without pain, then

presumably we can imagine a minimal physical/functional duplicate of the actual world but without both pain and elation; that we can imagine such a world but without pain, elation, and auditory experiences; and that we can imagine such a world but without pain, elation, auditory experiences, and visual experiences. Where do we stop? At what point in this series does the hypothesis become incoherent? Why can't we keep subtracting specific kinds of consciousness without end? But a zombie world is just the limit case in the series. So, if Marcus accepts that we can imagine a minimal physical/functional duplicate of the actual world but without pain, then it looks as though he should also accept the claim that we can imagine a zombie world. In any event, the specificity clause does not save the entailment principle.

Thus, EP, EP', and EP'' would seem to be vulnerable to rather straightforward counterexamples. If there are other restrictions on EP that Marcus intends, he does not indicate what they are. I conclude that his argument fails.

2.2 Can Marcus reject the entailment principle?

Marcus does not explicitly appeal to EP, EP', EP'', or anything of the kind. Therefore, he might respond that I have misconstrued his argument. But if that is so, then how does his argument work? He observes that there is no imaginable positive property (no subjective presence) that would enable us to imagine total subjective absence in the way that *being a shiny scalp* enables us to imagine the absence of hair. But without EP, or something very much like EP, that observation fails to undercut the conceivability of zombies. And it is plausible that we can imagine the absence of consciousness directly, without imagining a

positive property that entails such an absence. If nothing in the vicinity of EP grounds his denial that zombies are conceivable, then it is unclear what does.

However, Marcus does give one argument that might seem relevant. He writes,

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. (Marcus 2004: 483)

This is suspicious. How does it follow from (i) it is a subjective fact about zombies that they lack consciousness and (ii) there is nothing it's like to be totally subjectively absent that (iii) there is nothing of the what-it's-like variety we can use to imagine total subjective absence? The answer might seem obvious. If, in trying to imagine total subjective absence, we use something of the what-it's-like variety—a phenomenal quality—then we are not imagining *total* subjective absence: the phenomenal quality we use is subjectively present.

But that reasoning is confused. We use phenomenal qualities only to form a conception of a zombie world. The qualities are not themselves part of the hypothesis to which the conception corresponds. The confusion may stem from the idea of *using* a phenomenal quality in imagining. Does so using a phenomenal quality mean merely considering it? Or does it mean that the quality is instantiated in the imagined situation?

Take phenomenal redness. How might we use this quality in trying to imagine zombies? We do not imagine creatures that have experiences with that phenomenal

quality. Plainly, such creatures would not be zombies. Rather, we imagine a situation in which nothing has that (or any other) phenomenal quality. It is only in that sense that we use the quality. The idea that the conjunction of (i) and (ii) straightforwardly entails (iii) seems to assume that every quality we consider in imagining a situation S is part of S. But that assumption is mistaken. In imagining a pain-free world, we consider pain. But in the world we imagine, there is no such quality.

3. Is the zombie intuition based on a scope confusion?

Marcus recognizes that his view raises the question of why it seems that zombies are imaginable if in the end they are not. Here is his answer:

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 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. (Marcus 2004: 477-78)

Thus, Marcus would have us believe that the source of the zombie intuition is an elementary scope fallacy, akin to confusing *imagining a world in which Bush exists but Karl Rove does not* with *imagining a world in which Bush exists while refraining from imagining that Karl Rove exists*. This diagnosis strains credibility. We do sometimes

crudely misdescribe our modal intuitions with disastrous results; witness pre-Kripkean claims purporting to establish contingent identities. But when Kripke distinguishes *imagining Hesperus without itself* from *imagining Hesperus' not being the morning's last visible heavenly body*, we lose the inclination to claim that we can do the former. We recognize that our original description was inaccurate. By contrast, *imagining a zombie world* seems unproblematic even after we distinguish it from *imagining a minimal physical/functional duplicate world while refraining from imagining the instantiation of any state of consciousness*. Here there is little temptation to conclude that what we were doing all along was the proposed alternative. Even after careful reflection on Marcus' distinction, the zombie hypothesis seems as coherent as ever.

4. How not to answer the zombie challenge

I have argued that Marcus' attempts to undercut the claim that zombies are conceivable fail. His argument against the possibility of imagining total subjective absence relies on a dubious assumption about the nature and limits of imagination, such as EP. And his suggestion that the apparent conceivability of zombies derives from a scope fallacy lacks plausibility.

I mentioned at the outset that Marcus' strategy has the right form. This is particularly true of his argument against the possibility of imagining total subjective absence. To explain why, it will be useful to briefly consider another widely discussed anti-physicalist argument: Frank Jackson's (1982) knowledge argument. In Jackson's argument, Mary is a super-scientist who is raised in a black-and-white room and learns the complete physical truth from watching black-and-white science lectures. Then she

then leaves the room and sees colors for the first time. Intuitively, it seems that when she leaves she learns new phenomenal truths. The case is meant to support the claim that (roughly put) there are phenomenal truths that are not a priori deducible from the complete physical truth; and the latter claim is then used to support the anti-physicalist conclusion that there are phenomenal truths that are not metaphysically necessitated by the complete physical truth.⁷

Many aspects of the knowledge argument have been questioned.⁸ But most critics recognize that it presents a serious challenge to physicalism and that refuting it would require a substantive counter-argument. Benj Hellie (2004) puts the point nicely. He writes, “it won’t do to go out in search of some ‘fallacy,’ because there is no fallacy in the neighborhood” (Hellie 2004: 352).⁹

The same moral applies to the argument from the conceivability of zombies—and in particular to the claim that zombies are conceivable. Careful reflection on this conceivability claim seems only to confirm its plausibility. Philosophers who endorse it, such as Chalmers, commit no fallacy; they make no relatively simple error in reasoning. Consequently, suggestions that such a blunder underlies the claim are bound to fall flat.

To undermine the conceivability claim, critics must do more than show how some confusion might mislead us. They must expose a hidden incoherence in the zombie-world hypothesis. This is precisely what Marcus tries to do in arguing against the possibility of imagining total subjective absence. Although his argument fails, it provides a good model.¹⁰ Whether a successful argument for the same conclusion can be given remains to be seen.¹¹

References

- Alter, T. 2007. Does Representationalism Undermine the Knowledge Argument? In *Phenomenal Knowledge and Phenomenal Concepts: New Essays on Consciousness and Physicalism*, T. Alter and S. Walter (eds.). New York: Oxford University Press, pp. 65-76.
- Alter, T. Forthcoming. The Knowledge Argument. In the *Blackwell Companion to Consciousness*, S. Schneider and M. Velmans (eds.). Oxford: Blackwell.
- Ashwell, L. 2003. *Conceivability and Modal Error*. MA Thesis, University of Auckland.
- Balog, Katalin. 1999. Conceivability, Possibility, and the Mind-Body Problem. *Philosophical Review* 108: 497-528.
- Beaton, M. 2005. What RoboDennett Still Doesn't Know. *Journal of Consciousness Studies*, 12: 3-25.
- Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.
- Chalmers, David. 2002. Does Conceivability Entail Possibility? In *Conceivability and Possibility*, T. Gendler and J. Hawthorne (eds.). New York: Oxford University Press, pp. 145-200.
- Chalmers, David. 2003. The Content and Epistemology of Phenomenal Belief. In *Consciousness: New Philosophical Perspectives*, Q. Smith and A. Jokic (eds.). New York: Oxford University Press, pp. 220-72.
- Chalmers, D. J. 2004. Phenomenal Concepts and the Knowledge Argument. In *There's Something about Mary: Essays on Phenomenal Consciousness and Frank*

- Jackson's Knowledge Argument*, P. Ludlow, D. Stoljar, and Y. Nagasawa (eds.).
Cambridge: MIT Press, pp. 269–98.
- Churchland, P. 1989. Knowing Qualia: A Reply to Jackson. From his *A Neurocomputational Perspective*. Cambridge, MA: MIT Press, pp. 67-76.
- Hellie, Benj 2004. Inexpressible Truths and the Allure of the Knowledge Argument. In *There's Something about Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, P. Ludlow, Y. Nagasawa, and D. Stoljar (eds.).
Cambridge, MA: MIT Press, pp. 333-64.
- Jackson, Frank 1982. Epiphenomenal Qualia. *Philosophical Quarterly* 32: 127-36.
- Jackson, Frank 2003. Mind and Illusion. In *Minds and Persons: Royal Institute of Philosophy Supplement 53*. A. O'Hear (ed.). Cambridge: Cambridge University Press, pp. 251-71.
- Jackson, F. 2007. The Knowledge Argument, Diaphanousness, Representationalism. In *Phenomenal Knowledge and Phenomenal Concepts: New Essays on Consciousness and Physicalism*, T. Alter and S. Walter (eds.). New York: Oxford University Press, pp. 52-64.
- Kripke, Saul 1972. Naming and Necessity. In *The Semantics of Natural Language*, G. Harman and D. Davidson (eds.). Dordrecht: Reidel. Reprinted as *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.
- Marcus, Eric 2004. Why Zombies are Inconceivable. *Australasian Journal of Philosophy* 82: 477-90.
- Farid Masrour. N.d. Are Zombies Positively Conceivable? University of Arizona manuscript.

Shoemaker, Sidney. 1975. Functionalism and Qualia. *Philosophical Studies* 27: 291-315.

¹ Throughout the paper I go back and forth between the claim that zombies are conceivable and the claim that a zombie world is conceivable, where a zombie world is a minimal physical and functional duplicate of the actual world but without consciousness. The differences between these claims do not matter much here.

² This virtue is lacking in other recent discussions of the argument. See, for example, Beaton 2005, p. 21.

³ Lauren Ashwell (2003) and Farid Masrour (n.d.) have disputed the conceivability of zombies on similar grounds. My criticisms of Marcus' arguments apply to Ashwell's and Masour's *mutatis mutandis*.

⁴ I will ignore various subtleties, such as Chalmers' treatment of vague hypotheses. Here further precision is not needed. Whenever I speak of conceiving, I have in mind ideal, primary, positive conceiving, unless otherwise specified.

⁵ Of course, we know the actual world contains consciousness. But then we also know that water is H₂O. Such knowledge is a posteriori and therefore irrelevant to primary conceivability.

⁶ I thank David Chalmers for helpful suggestions about EP and its problems.

⁷ This formulation of the knowledge argument is imprecise. For example, the deduction base mentioned in the premise about non-deducibility should include all indexical information, such as the truths expressed by "I am s" and "Now is t" where s and t are descriptions specifying unique subjects and times. For a precise formulation (and forceful defense) of the argument, see Chalmers 2004. Jackson (2003, 2007) now rejects the argument. I criticize his basis for so doing in Alter 2007.

⁸ For a survey, see Alter forthcoming.

⁹ *Pace* Paul Churchland (1989).

¹⁰ Another good example is Sidney Shoemaker's (1975) argument from self-knowledge. Shoemaker argues that the conceivability of zombies ("absent qualia") cannot be reconciled with plausible claims about first-person knowledge of our own experiences, because zombies would have the same introspective mechanisms that we have. Katalin Balog (1999) gives a related argument. I think these arguments, like Marcus', fail. For criticisms of Shoemaker's argument, see Chalmers 1996, pp. 192-96. For criticisms of Balog's argument, see Chalmers 2003, section 4.3, especially fn. 19 on p. 258.

¹¹ I presented this paper as part of a talk I gave at the *Physicalism and Beyond* workshop at the 2006 GAP.6 conference in Berlin. I thank the participants for a useful discussion. For helpful comments, I thank David Chalmers, Eric Marcus, and an anonymous referee for *Disputatio*. This paper benefited from my participation in John Heil's 2006 NEH Seminar, *Mind and Metaphysics*.