

# All You Zombies.

## David Chalmers' Metaphysical Solipsism

Steffen Borge

---

David Chalmers' doctrine that Zombie-worlds are conceivable and thus possible has as a consequence, if sound, that materialism is false. Chalmers' conceivability argument, I will argue, is best understood as a sceptical challenge to materialism, which, if accepted, yields metaphysical solipsism.

Chalmers resists metaphysical solipsism by appeal to epistemological considerations about other minds and inference to the best explanation. This, however, generates a dilemma for Chalmers, as I shall argue. If we are justified in believing in other minds in this world, then we are justified in all relevant possible worlds, and then there is no reason to postulate mental stuff over and above the physical. If we are not justified, then by the same token, metaphysical solipsism follows. Thus, the challenge against materialism fails if we are not willing to endorse metaphysical solipsism, and we have yet to see any convincing arguments for the latter view.

### 1. The Conceivability of Zombie-Worlds and the Refutation of Materialism

Let me first present Chalmers' argument. The main premise for Chalmers' argument is that there is consciousness in our world (Chalmers, p. xii-xiii, 1996). According to Chalmers we must take seriously the what-is-it-like phenomenological feature of consciousness if we are to explain, and not explain away, consciousness.

It might be the case that consciousness supervenes on the physical in our world, but we can imagine things being otherwise (Chalmers, p. 37-38). Chalmers is careful to avoid standard Kripkean objections from *a posteriori* necessity (Kripke, 1996). *Prima facie* it seems conceivable that water is not H<sub>2</sub>O; on the other hand, given Kripkean considerations, and assuming that "water" and "H<sub>2</sub>O" are rigid designators, the statement "Water is H<sub>2</sub>O" expresses a necessary truth. However, if we knew more about the relation between the mental and the physical could it not turn out that the two were identical in the same way water and H<sub>2</sub>O turned out to be?

Following Davies and Humberstone (Davis and Humberstone, 1980) Chalmers distinguishes between the primary and secondary intensions of terms (Chalmers, p. 57). The primary intension of the term "water" picks out watery stuff. In this case, we say that the term's intension is the ordinary prescientific conception of water. Both

Putnam's earth water and twin-earth water will be picked out by the term's primary intension (Putnam, 1984). When we say that we can imagine water not being H<sub>2</sub>O, we mean to say that we can conceive of watery stuff that has another atomic structure. The secondary intension of a term, on the other hand, only comes into play when the reference of the term is fixed for some linguistic community in relation to some object or natural kind. With the secondary intension we intend to refer to not only watery stuff as such, but rather "*This or that* kind of watery stuff". And then the empirical facts about these objects or kinds limit the possible counterfactual situations in which the term can be correctly applied.

There is, however, no applicable secondary intension to be had in relation to the term "consciousness", since anything that satisfies the what-is-it-like requirement of consciousness, in fact, counts as consciousness. Thus conscious experiences could be instantiated in physical systems radically different from ours (Chalmers, p. 97). And more importantly, we could then imagine somebody with the same physical constitution as ourselves who did not instantiate the same conscious experiences.

The argument against materialism then goes as follows:

P1. Every conceivable world is logically possible (Chalmers, p. 66).

P2. I can conceive of a world that is a molecule for molecule physical replica of our world, in which there is no consciousness thus such a world is conceivable. Call such a world a Zombie-world (Chalmers, p. 123).

From P1 and P2 it follows that:

C1. Zombie-worlds are logically possible.

P3. If Zombie-worlds are logically possible, then not all facts about consciousness are exhausted by physical facts (Chalmers, p. 123).

By modus ponens it follows from C1 and P3 that:

C2. All facts about consciousness are not exhausted by physical facts.

P4. If materialism is true, then all facts about consciousness are exhausted by physical facts (Chalmers, p. 41).

By modus tollens it follows from C2 and P4 that:

C3. Materialism is false.

So there is more to conscious experiences than is dreamt of in a materialistic metaphy-

Steffen Borge

sics.

Certain doubts remain, however. Exactly how are we to come by these all-important Zombie-worlds? That I do not inhabit a Zombie-world is clear from my own conscious experiences. An infallible Cartesian certainty prevents me from entertaining doubts about my own consciousness (Chalmers, p. 183). I know that I am not a Zombie, but what can I say about you in light of the logical possibility that I might find myself in a partial Zombie-world?

*a partial zombie world: one in which I am at the center, conscious with all the relevant concepts, but in which some other people are zombies (Chalmers, p. 133).*

Chalmers insists that there is no reason, in our world, for not believing in other minds. By noting the strong regularities between conscious experiences and physical or functional states in our own case, we can postulate psychophysical laws and infer from them to the consciousness of other beings like ourselves (Chalmers, p. 246). This provides us with a reasonable justification for believing in other minds. This epistemological argument, however, will either take the wind of Chalmers' Zombie-world example or be effectively undermined by it.

## 2. Inference to the Best Explanation or Metaphysical Solipsism

Chalmers is untroubled by the explanatory gap opened by his Zombie-scenario.

Now my Zombie twin is only a logical possibility, not empirical one, and we should not get too worried about odd things that happen in logically possible worlds (Chalmers, p. 180).

This untroubled attitude is misguided. If Chalmers' argument against materialism is to preserve its bite, then we have every reason to worry about Zombies in our own vicinity.

Chalmers' psychophysical laws will require nothing over and above the physical (Chalmers, p. 217). Materialism would then, *mutatis mutandis*, be the best explanation for other minds in all the relevant possible worlds, since possible conscious experiences in other beings are, according to Chalmers, in principle hidden from us. Let us quantify over the set of possible worlds with the same physical constitution as ours. Domain  $D: \{W_{\text{Actual}}, W_{\text{Replica world 1}}, W_{R2}, \dots, W_{RN}\}$ . Let  $P$  be the property of having the same physical constitution as our world. Let  $J$  be the property of being justified in believing in other minds. In order to avoid scepticism about other minds in our world, Chalmers argues that we are justified in postulating psychophysical laws that relate the mental to

the physical in a systematic way. Let  $L$  be the psychophysical laws. On the basis of  $P$  and  $L$ , we get  $J$  in our world, but we can exploit this in order to get  $J$  in any world in  $D$ . Take any arbitrary world  $W$  in  $D$ ; if we get  $J$  in  $W$  we get  $J$  in all worlds in  $D$  by universal generalisation.

Accordingly, we can know *a priori* that if we are justified in one world in believing that we are not in a partial Zombie-world or "looking" at a Zombie-world, then there is no world where we would be justified in having Zombie-qualms. We cannot conceive of a world in  $D$  where we would be justified in believing that there are Zombies of the required kind.

Inference to the best explanation then leads us straight back to a *ceteris paribus* endorsement of materialism. Granted that inference to the best explanation is a non-deductive inference pattern that does not give us infallibility, so the brute logical possibility of Zombie-worlds remains. On the other hand, we have no epistemic access to this possibility, and it is ruled out as an unjustified possibility to consider if we are justified in one world in believing in other minds. There seems to be an artificial distinction about standards of justification in our world as opposed to cross-world considerations. In our world, according to Chalmers, we have good reasons for postulating simple psychophysical laws and we need not worry about other possible worlds. In a one-world perspective of the actual world we can say that consciousness arises from the physical and/or that consciousness is nothing above the physical. This would be the best and simplest explanation for the before-mentioned regularities. The actual does not licence any postulation of something over and above physical and functional properties. So without cross-world considerations, inference to the best explanation favours materialism, since this is the only way, as Chalmers puts it, things fall into place (Chalmers, p. 246). If we have an epistemic duty to believe in the best available explanation, then we are obliged to believe in other minds in all relevant possible worlds on the basis of the inference to the best explanation as given by Chalmers. But not only that, we are epistemologically obliged to believe in materialism in all the relevant logically possible worlds.

I have, however, not proven materialism. Only on the assumption that scepticism about other minds is not justified have we been able to vindicate materialism as the best explanation in all relevant possible worlds. But is there not a difference between claiming that (I) in no worlds of type  $T$  are we justified in doubting other minds and (II) that we are justified in believing that there are no worlds of type  $T$  in which there are no other minds? We might have good reasons to believe that very misleading circumstances are possible. Let any such circumstances be the absence of other minds, then (I) is true and (II) is false. This objection, however, fails to take into account the fact that the various worlds in (I) and (II) are duplicate worlds. How could the same evidence in one world justify our beliefs about other minds, but not justify them in duplicate

worlds? Given that we can make sense of duplicate worlds and that the mental is hidden from physical sciences, then (I) and (II) do collapse into each other, even though (I) does not imply (II).

It is important to see that this case where (I) and (II) collapse into each other is disanalogous to the lottery paradox. Imagine a set of lottery worlds and let that set be all possible outcomes of an imaginary lottery. (I\*) In all of the worlds we are justified in believing that you will not win the lottery. (II\*) We are justified in believing that you will not win the lottery in all of the worlds. (II\*) is false, since there is some world among the lottery worlds where you will win, even though we are not justified in that world, as in the other worlds, in believing that you will win. What makes the difference here is that in the lottery case there is an observable fact of the matter about who wins and who loses. This is disanalogous to the Zombie-case where there is no observable fact of the matter to be found at the end of inquiry. There is a discernible answer to the question whether you win or not, but not to the question whether you are a Zombie or not. What distinguishes the lottery worlds are the observable outcomes of the lottery. The worlds in the set of lottery worlds are not duplicate worlds, unlike the Zombie-worlds, which must be seen as duplicate worlds if they are to present a challenge to materialism.

A critic, however, might object that these epistemological considerations do not get to the heart of the matter; the point is that we could be wrong. Your best explanation of people's behaviour might be that they have conscious experiences, but that does not ensure you that you get it right. Your psychophysical laws are perhaps your best means for understanding and predicting behaviour, but there is more to consciousness than being explainable in terms of psychophysical laws. Given the conceivability of Zombie worlds, there will in logical space be at least one possible world where our psychophysical laws fail to pick out connections between the physical and the mental, since there are no mental properties in that world. Furthermore, psychophysical laws do not give us conclusive evidence for other minds, so we are not justified in believing in other minds on the basis of these. It follows that if justification breaks down for one of the worlds in *D*, it breaks down in all of them. If we pick an arbitrary world in *D* and conclude that we do not get *J* there, then by universal generalisation we will not get *J* in any world in *D*.

In this scenario, where exclusive emphasis is given to conclusive evidence in justification of beliefs about other minds, Chalmers' Zombie-world argument retains its bite against materialism. The price to pay is the vindication of scepticism about other minds and consequently of metaphysical solipsism. In this reading the Zombie-world argument amounts to little more than a sceptical argument against materialism. I have certainty about my own consciousness, but as for the rest of you, you might all be Zombies. As with all sceptical arguments, there is a question of burden of proof.

### 3. Scepticism and the Burden of Proof

It is not my task in this paper to disprove scepticism, but it is worth while having a short look at it. Consider the following argument for scepticism about other minds. Imagine that a demon has arranged it such that we all live in partial Zombie-worlds. This is clearly conceivable, thus logically possible. If we cannot prove or show the Zombie-hypothesis to be false, then we do not know it to be false and that in effect casts doubt upon our beliefs about other minds. The net effect of this scepticism is that for any belief  $p$  (believing in other minds), we do not know that  $p$ , since we cannot with certainty, due to the Zombie-hypothesis, know that not- $p$  is false (Lehrer, 1971). We are left with metaphysical solipsism.

Again and again Chalmers argues that the burden of proof is on the materialist opposition to show that there is something wrong with the conceivability of Zombie-worlds. But in view of scepticism about other minds, it is time to shift the burden of proof. Why demand that one must prove or show that the contrary to a belief that  $p$  is false in order for one to be justified in believing that  $p$ ? We might not be able to show the Zombie-world hypothesis to be false or that Chalmers misdescribes Zombie-worlds, but it still might be false. Nor does it follow from this that we cannot show the Zombie-world hypothesis to be unjustified. Even though we are presently incapable of proving the Zombie-world hypothesis to be false, we might still regard it as unjustified, since there are no reasons to believe it in any of the relevant worlds.

If, indeed, Chalmers' position ends up in metaphysical solipsism when taken as a real challenge to materialism, it is reasonable to shift the burden of proof onto the sceptic. Why should I doubt other minds? We need something more than brute logical possibility in order for us to embrace Chalmers' metaphysical solipsism. We need some positive arguments for the reasonableness of that possibility in the realm of the relevant possible worlds. I cannot see that Chalmers has given us that, nor can I see how it can be got from Zombie-world considerations.

### References

- Chalmers, D. *The Conscious Mind*. Oxford University Press, 1996.
- Davies, M. and Humberstone, L. "Two Notions of Necessity", *Philosophical Studies*, 38, 1980.
- Kripke, S.A. *Naming and Necessity*. Harvard University Press, 1996.
- Lehrer, K. "Why Not Scepticism?". *The Philosophical Forum*, 2.3, 1971.
- Putnam, H "The Meaning of 'Meaning'". In *Mind, Language and Reality*, Cambridge University Press, 1984.