

Karol Polcyn

Conceivability, Possibility and Rationality

The zombie argument recently developed by David Chalmers (1996, 2010) aims to establish that zombie worlds (worlds identical to our world in all physical respects but lacking consciousness) are possible and that, therefore, materialism is false. The key move in this argument is that zombie worlds are possible because they are conceivable. While Chalmers gives a few different reasons in support of this inference, perhaps the most intuitive one is that there are no examples of unreliable modal intuitions outside of the mind-body domain and therefore *it would be ad hoc* to deny the possibility of zombies. Here I argue that the denial of the possibility of zombies is not ad hoc. There is a certain crucial difference between the zombie intuition and other modal intuitions, which explains that while other modal intuitions are reliable, we have no reason to trust the zombie intuition. Roughly, the difference is that while other modal intuitions have some *rational explanation*, the zombie intuition does not. The only explanation of the zombie intuition we can think of is purely conceptual; that is, zombies are conceivable only because of some conceptual differences between phenomenal and physical concepts.

Consider first the paradigm cases of conceivability intuitions, ones that Chalmers takes to be reliable guides to possibility. Those are intuitions about the possible falsity of standard theoretical identities, such as the intuition that water might not have been H₂O or that heat might not have been molecular motion. As Chalmers observes, following Saul Kripke (1980), it is not, strictly speaking, conceivable that water might not have been H₂O but only that water might not be H₂O. In other words, water without H₂O is conceivable *primarily* (across possible actual worlds) and not *secondarily* (across possible counterfactual worlds). Let's focus then on the primary conceivability of worlds in which there is water but no H₂O. According to Chalmers, it is conceivable that water might not be H₂O because it is conceivable that watery

stuff (the stuff that manifests certain observable properties, such as transparency, liquidity, drinkability, etc.) might not be H₂O. The conceivability of the latter gives rise to the conceivability of the former because watery properties are the properties that fix the reference of 'water' (those are the properties that determine what 'water' refers to in any possible world considered as actual). The conceivability of worlds in which water is not H₂O is thus ultimately grounded in the way we use the word 'water'. So, in this sense, Chalmers (2002a) assumes that this conceivability is simply a matter of conceptual coherence.

But the crucial question is why it is conceivable that watery stuff might not be H₂O in the first place. Chalmers never addresses this question. He might say that this intuition is conceptually coherent too. But there is certainly more to say about this intuition. It seems quite clear that this intuition has some rational explanation. Roughly, the explanation of its rationality has to do with the fact that whereas watery properties (higher-level properties) depend nomologically on the property of being H₂O (lower-level property), the property of being H₂O does not depend nomologically on watery properties. Since the property of being H₂O does not depend nomologically on watery properties, we find it natural to think that watery stuff might not be H₂O. By contrast, we assume that H₂O without watery properties is impossible given that watery properties depend nomologically on the property of being H₂O, and this explains why we cannot conceive of worlds in which there is H₂O but no watery properties (Chalmers, Jackson 2001). Notice that the assumptions about nomological relations between watery properties and H₂O are not themselves established on the basis of what we find conceivable; rather, we know they are true on the basis of our overall theory of the world.

Consider now the primary conceivability of worlds in which heat is not molecular motion (the worlds in which there is heat but no molecular motion and ones in which there is molecular motion but no heat). The primary conceivability of such worlds is, of course, grounded in the conceivability of worlds in which there is the sensation of heat which is not caused by molecular motion (but some other phenomenon) and those in which there is molecular motion which does not cause the sensation of heat (but some other kind of sensation). Chalmers is right to assume that if such worlds are conceivable, then given that the property of causing heat sensations is the property that fixes the reference of 'heat', it is conceivable that heat as such and molecular motion might not be identical in another actual world.

But again, there is the further question of what explains the conceivability of worlds in which the causal relation between the sensation of heat and molecular motion is broken. It seems that the answer has to be that the relation of causation, in particular the relation of causation between heat and heat sensations, is contingent in the sense of being dependent on certain contingent circumstances. In particular, whether or not heat causes the sensation of heat depends on the perceptual apparatus of creatures which are exposed to heat. As a matter of fact, we are so constituted that we experience heat in the presence of heat. But we can certainly imagine that our

planet might be inhabited by creatures which react to heat quite differently. Those other creatures do not have the sensation of heat in the presence of heat but rather some other kind of sensation. Accordingly, they experience heat not in the presence of heat but some other physical phenomenon (Kripke 1980: 131–132).

To sum up, in all the paradigm cases of conceivability there is some rational explanation of why we find conceivable what we do. Chalmers assumes that in all those cases conceivability is simply a matter of conceptual coherence. My point is that this does not fully characterize those cases. The conceivability of worlds in which water is not H₂O or worlds in which heat is not molecular motion is not only conceptually coherent but also has some rational explanation.

Consider another case illustrating the key point. A mile high skyscraper is certainly conceivable even though no one has actually built one (Chalmers 1996: 36–37). There is nothing conceptually incoherent about this intuition. But it is clear that in addition to being conceptually coherent this intuition has rational explanation; the explanation is that a mile high skyscraper is not impossible nomologically, which is something that we know on a posteriori grounds. In the paradigm cases discussed above, the explanation of our conceivability intuitions did not appeal to nomological possibility but to some other intuitions: the intuition that there is no nomological dependence of lower-level properties on higher-level properties and the intuition that causation is contingent. Still, to the extent that the paradigm cases have some rational explanation, they resemble the case of the skyscraper.

Notice also that in all the paradigm cases of conceivability intuitions we are assuming that the relevant properties are *not identical*: the relation of causation (as between heat and the sensation of heat) is a relation that holds between different properties and the relation between watery properties (higher-level properties) and the property of being H₂O (lower-level property) is not the relation of identity, either. It seems reasonable to say that this is what ultimately explains why we find it conceivable that heat might not be molecular motion or that water might not be H₂O.

If we now compare the paradigm cases of conceivability with the zombie intuition, we can easily see that the zombie intuition is fundamentally different. There is no corresponding rational explanation of the zombie intuition. To begin with, we cannot assume that phenomenal and physical properties are not identical in order to explain the conceivability of zombies. This is because we have strong empirical reasons to think that phenomenal and physical properties are identical.

More specifically, the difference between the zombie case and other cases of conceivability can be described as follows. Unlike in the case of watery properties and H₂O, we have no reason to think that consciousness does not depend nomologically on physical properties. So we cannot explain the conceivability of zombies by assuming that consciousness does not depend nomologically on physical properties.

Secondly, the conceivability of zombies cannot be explained by assuming that in zombie worlds physical states do not cause conscious states and that this is so because the relation of causation is contingent. The trouble with this suggestion is that

it is not rational to think that in zombie worlds physical states do not cause conscious states since, by assumption, the sensory apparatus in zombies functions in just the same way as it does in our case. There is also the further question of why physical states in zombie worlds do not cause conscious states at all. When we assume that the relation of causation between heat and heat sensations is contingent, we are assuming that heat always causes some sensations given the presence of the appropriate sensory apparatus so that, even if it does not cause the sensations of heat, it causes some other kind of sensations. But nothing like this happens in the zombie case. In the zombie world, the causation relation is broken in a much stronger sense since no conscious states of whatever kind are caused in those worlds.

Finally, the conceivability of zombies could not be explained by assuming that it follows from the conceivability of different laws of nature. This is because the conceivability of zombies does not presuppose that the laws of nature are different from what they actually are. This is true at least about physical laws. Zombie worlds are conceivable even if we suppose that physical laws remain the same.

Probably, there are also some nonphysical laws, namely laws that link consciousness understood as a nonphysical property with physical processes. If consciousness depends on physical processes in virtue of such nonphysical laws (psychophysical laws, as Chalmers calls them), then it is not conceivable that there should be worlds physically identical to our world and obeying psychophysical laws but lacking consciousness. However, we cannot explain the conceivability of zombies by assuming that it involves the conceivability of worlds in which psychophysical laws are different than in our world. This is because we cannot assume there are any psychophysical laws, unless we already believe that zombies are possible and, obviously, the belief that zombies are possible can hardly be regarded as part of our commonsense rational thinking informed by science.

So it appears that the zombie intuition has no rational explanation. This should be no surprise once we realize that the only explanation of the zombie intuition is purely conceptual: the intuition cannot be ruled out even on ideal rational reflection due to the fact that phenomenal concepts are distinct from physical concepts so that phenomenal truths are not implied a priori by physical truths. While there are various ways to explain why phenomenal and physical concepts are distinct concepts, Chalmers (2002b) himself emphasizes that phenomenal concepts, unlike physical concepts, are not structural but qualitative concepts, which means that they represent their referents not in terms of structural properties but qualitative ones.

I take it that the fact that the zombie intuition has no rational explanation undermines Chalmers' argument for the possibility of zombies. Since the paradigm cases of conceivability have some rational explanation and the zombie intuition does not, it is *not ad hoc* to deny that the conceivability of zombie entails possibility.

It should be pointed out that Chalmers does not justify the inference from the conceivability to the possibility of zombies only by assuming that the denial of this inference would be *ad hoc*. In addition, he argues that there is an intrinsic connection

between ideal conceivability and possibility based on the rational roots of our modal concepts. Ideally conceivable worlds must be metaphysically possible because those worlds have an important explanatory role to play (they account for various aspects of our language and thought), and the denial of their possibility has no explanatory role. That is, in fact, the most fundamental reason why conceivability entails possibility, according to Chalmers.

In response, I agree that the explanatory role of ideally conceivable (logically possible) worlds is one reason why we speak of those worlds as possible worlds at all. But my point is that we have additional criteria for evaluating whether or not logically possible worlds are possible. As it turns out, zombie worlds are those ideally conceivable worlds that do not satisfy those additional criteria, and this gives us the reason to doubt the real possibility of zombies. It does not really matter that the denial of the possibility of zombies does not explain any modal data.*

REFERENCES

- Chalmers D. (1996), *The Conscious Mind*, New York: Oxford University Press.
- Chalmers D. (2002a), *Does Conceivability Entail Possibility?* [in:] *Conceivability and Possibility*, T. Gendler, J. Hawthorne (eds.), New York: Oxford University Press, 145–200.
- Chalmers D. (2002b), *Consciousness and Its Place in Nature* [in:] *Philosophy of Mind. Classical and Contemporary Readings*, D. Chalmers (ed.), New York: Oxford University Press, 247–272.
- Chalmers D. (2010), *The Two-Dimensional Argument against Materialism* [in:] *The Character of Consciousness*, ed. D. Chalmers, New York: Oxford University Press, 141–206.
- Chalmers D., Jackson F. (2001), *Conceptual Analysis and Reductive Explanation*, „Philosophical Review” 110, 315–361.
- Kripke S. (1980), *Naming and Necessity*, Cambridge (MA): Harvard University Press. First published in *Semantics of Natural Languages*, D. Davidson, G. Harman (eds.), Dordrecht: Reidel, 1972.

* This article is a part of a project sponsored by National Science Centre (decision nr DEC-2011/01/B/HS1/05416).