

## *Zombies, duplicates, and theories of consciousness (Commentary on Fischer and Sytsma)*

Keith Frankish

---

A philosophical zombie is an imaginary being that is a molecular level ('micro-physical') duplicate of a fully conscious human but lacks a certain kind of consciousness. The relevant kind of consciousness is not easy to describe but is taken to be introspectively obvious—it is the 'subjective feel' of experience. David Chalmers likens it to an inner light; without it, 'all is dark inside' (Chalmers, 1996, p. 96). If zombies are clearly conceivable, then, it seems, this kind of consciousness is not physical. A microphysical duplicate of a conscious human will share all the original's physical properties—where 'physical' includes both properties described by physics and properties ultimately realized in those properties. If we can clearly conceive of all these properties being present without this kind of consciousness, then the latter is not itself a physical property. This is the zombie argument against physicalism.

So, *can* we conceive of zombies? Can we clearly imagine a being that shares all the physical properties of a human but is 'all dark' inside? As Fischer and Sytsma show, how people answer depends on the label we use for the beings in question. If we call them *zombies*, as philosophers usually do, then people are significantly more likely to say they are conceivable than if we call them simply *duplicates*. The label strongly influences people's conceivability judgements, and this fact undercuts the evidence for the conceivability of philosophical zombies (Fischer and Sytsma, 2021). This isn't surprising. For most people, 'zombie' conjures up an image of a horror movie monster, one of the living dead, who feels nothing. (It's OK to shoot zombies and beat them up with cricket bats.) Of course, *philosophical* zombies are supposed to be very different from these monsters, but the term 'zombie' biases us to imagine them as creatures with no inner life. 'Duplicate' lacks this biasing effect and promotes judgements in line with a physicalist view of consciousness. The upshot is that the zombie argument fails.

Personally, I find this result congenial. I'm a physicalist, and I don't believe that the zombie argument tells us anything about the nature of human consciousness. Indeed, I would be glad to see the word 'zombie' expunged from debates about consciousness. Since I don't believe in the existence of the kind of consciousness that zombies are supposed to lack, there is a sense in which I think *we* are zombies—which suggests to some that I hold the ridiculous view that we have no feelings and are incapable of suffering (e.g., Strawson, 2019).

Getting rid of the term ‘zombie’ should help to prevent this misinterpretation.

Despite this, I don’t think that Fischer and Sytsma’s study shows that people do not have the dualist intuition that the zombie argument relies on. For it may be that the term ‘duplicate’ also has a biasing effect. The key point is that in imagining a philosophical zombie it is not enough to imagine a physical duplicate. We must imagine a *bare* physical duplicate—a duplicate in all physical respects *and no others*. We must not imagine a *rich* duplicate, which has properties that normally accompany these physical properties in virtue of contingent natural laws, causal or otherwise. Physicalism is not the view that consciousness is caused by, or accompanies, the physical, but that it *is* physical. So, in trying to conceive of zombies we must be ready to imagine variations in the background laws of nature. The task is like imagining duplicate fire that does not burn or duplicate cyanide that is not poisonous. Now Fischer and Sytsma’s study does not make it clear that subjects must imagine a bare duplicate rather than a rich one, and it may be that participants are doing the latter. Anecdotally, laypeople often think of physicalism as the view that mental states are dependent on physical ones rather than identical with them. Moreover, it may be that the word ‘duplicate’ tends to evoke a rich reading. Just as ‘zombie’ may encourage participants to imagine the absence of something, obscuring their physicalist intuitions, so ‘duplicate’ may encourage them to imagine the presence of something, obscuring their dualist ones. At any rate, it is a hypothesis worth testing.

Let’s pull back a bit now and ask what is at stake here. Suppose ‘duplicate’ does tend to invite a rich reading and that when people are asked specifically if a *bare* physical duplicate would be dark inside, they respond that it would. Would that show that physicalism is false? No. It would show that we have, or can be induced to form, a conception of consciousness as something non-physical—a subjective glow that is produced by the physical processes. But it wouldn’t show that the kind of consciousness we actually possess is like that. To arrive at that conclusion, we would need the further premise that our conception of consciousness is accurate, perhaps because introspection fully reveals the nature of consciousness to us. And that claim is an empirical one, whose truth cannot be determined by reflection on our concepts. Similarly, of course, if people resist the idea that a bare duplicate might be dark inside that would not establish that consciousness is physical, only that we conceive of it as physical. In each case, the imaginative exercise would tell us something about human psychology or culture, relevant to what Chalmers has called the *meta-problem of consciousness* (the problem of explaining why we think consciousness poses unique explanatory problems; Chalmers, 2018), but it would shed little light on the nature of consciousness itself. The history of science is

a history of folk conceptions of natural phenomena being overthrown or radically revised.

This points to what the word ‘zombie’, with all its baggage, is really doing in the debate. It is not revealing the implications of our intuitive grasp of the nature of consciousness but encouraging us to form and endorse a particular *theory* of consciousness—a theory that treats consciousness as a psychic essence distinct from all functional processes. For this purpose, the term is well suited. The depiction of movie zombies encourages us to adopt—or rather, to suspend disbelief in—a certain theory of life. Zombie bodies perform many of the functions of life but are dead inside. It seems they lack some animating essence—an essence that can be drained from a living body by a zombie’s attack. Philosophical zombies are a psychological parallel. Their brains perform all the functions of consciousness but are not conscious inside. They lack a phenomenal essence. If movie zombies are the living dead, then philosophical ones are the conscious unconscious, and the theory of consciousness implicit in the latter notion may be no better than the theory of life implicit in the former.

It is tempting to think that we have an immediate grasp of what consciousness is and that thought experiments can reveal important truths about its nature. But as Fischer and Sytsma show, our judgements in these experiments can be influenced by irrelevant linguistic factors. It is more likely that such thought experiments are tacitly encouraging us to adopt specific theories of consciousness. There is nothing wrong with that, provided we realize that that is what they are doing and that the theories they promote may be bad ones.

## References

- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, D. J. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies*, 25 (9–10): 6–61.
- Fischer, E. and Sytsma, J. (2021). Zombie intuitions. *Cognition*, 215.
- Strawson, G. (2019). A hundred years of consciousness: “a long training in absurdity”. *Estudios de Filosofía*, 59 (January–June): 9–43.

*The University of Sheffield*  
*k.frankish@gmail.com*