

Science vs. Scientism in Consciousness Research

A Reply to Ann-Sophie Barwich

Philip Goff

To be published in a book symposium hosted by the Social Epistemology Review and Reply Collective (SERRC), with the subject of Moti Mizrahi (Ed.) (2022) *For and Against Scientism: Science, Methodology, and the Future of Philosophy*, Rowman & Littlefield Publishers.

I am very grateful to Ann-Sophie Barwich for taking the time to comment on my work, in her paper '[Between electrical light switches and panpsychism: Scientism and the responsibilities of the humanities in the twenty-first century](#)' (unless otherwise stated, all citations are to this work). However, it's a little tricky to formulate a response given that Barwich doesn't discuss any of the arguments I give in support of my position. Usually when academic philosophers respond to my work, they discuss my academic writing, for example, my academic book *Consciousness and Fundamental Reality*, which is perhaps the most detailed version of my case against physicalism and in support of panpsychism. Instead Barwich has chosen to focus on my book *Galileo's Error: Foundations for a New Science of Consciousness*, which is a more simplified version of those arguments aimed at a general audience. But even in *Galileo's Error*, I give the knowledge argument and the zombie argument in support of the thesis that we cannot fully account for conscious experience in the terms of physical science. Barwich says absolutely nothing about these arguments. It's not obvious how to give a response when in substance no objection has been given.

Instead of dealing with my arguments, Barwich accuses me of 'science skepticism,' which she suggests is akin to vaccine skepticism and climate change denial. Whatever I'm doing, there's clearly a huge and crucial difference between my work and vaccine skepticism/climate change denial: I'm not disputing *any* claims supported by the scientific community. I do argue that science cannot fully account for consciousness. But that's because I don't think dealing with consciousness is a purely scientific task, and which tasks fall in the domain of science and which in the domain of philosophy is a contentious question about which it is reasonable to disagree.

Here's an example to illustrate this point. Sam Harris (2012) has notoriously argued that all ethical questions can be answered by science. Most people think that's false. Whilst our scientific knowledge is of course highly relevant to many ethical issues, it is not a purely empirical question whether, say, it's okay to kill chickens for food. Imagine someone who disputed this by saying, 'That's science skepticism! Whilst we have not yet designed the right experiment to show whether chickens have a right to life, the great success of science shows that we will in the end.' Whether or not this kind of pure scientism about ethics could be defended, it would be absurd to describe its opponents – the vast majority of philosophers – as 'science skeptics.'

To take another example, philosophers are split close to 50/50 on whether or not Platonism is true.¹ One of the most common arguments for Platonism is that Platonic entities are required to undergird

¹ <https://survey2020.philpeople.org/survey/results/all>

mathematical truth. Opponents of Platonism may defend a fictionalist or conventionalist account of mathematical truth. Now almost everyone agrees that this dispute between Platonists and their opponents is something we're going to settle with philosophy rather than observations and experiments. Imagine an objector saying 'That's science skepticism! We just haven't found the right experiment yet to test Platonism.' Again, whether or not this unusual application of methodological naturalism can be sustained, it's clearly ridiculous to accuse the vast majority of philosophers who think is a philosophical rather than a scientific questions of 'science skepticism' akin to climate change denial.

Likewise, the division of labour between science and philosophy in dealing with consciousness is itself a philosophically contentious issue. In my view, explaining consciousness is not a purely scientific question because consciousness is not a publicly observable datum. We can't look inside someone's head and see their feelings and experiences. We know that consciousness exists not from observation and experiment, but rather from our immediate, private awareness of our own feelings and experiences. Now, science is used to dealing with things we can't observe: fundamental particles, quantum wave functions, maybe even other universes. But there's an important difference in the case of consciousness. In all of these other cases, we theorise about things we cannot observe to explain what can be observed. In the unique case of consciousness, *the thing we are trying to explain* is not publicly observable.

That's not to say that philosophy *alone* can deal with consciousness. I think we need *both* science *and* philosophy to deal with consciousness (I imagine Barwich thinks so to, although no doubt we have a different view of the division of labour). One of the things the science of consciousness rightly aspires to put together is a systematic account of which kind of neural activity is necessary and sufficient for consciousness. But what we ultimately want from a theory of consciousness is not merely correlations but an explanation of *why* such and such brain activity is correlated with consciousness; because consciousness is not a publicly observable phenomenon, I don't think *this* question can be answered with an experiment.

What we bump up to at the point in enquiry is the mind-body problem, the philosophical challenge of bringing consciousness and physical reality together in a single, unified, worldview. Here are three possibilities:

1. *Materialism* – Physical reality is fundamental and consciousness emerges from physical processes in the brain.
2. *Panpsychism* – Certain facts about consciousness are fundamental and physical reality emerges from those underlying consciousness facts.
3. *Dualism* – Both consciousness and physical reality are equally fundamental but distinct.

These philosophical theories are empirically equivalent, at least in certain forms. For any empirical data concerning correlations between consciousness and brain activity, each theory will simply account for that data on their own terms. Suppose, for example, we discover empirically that consciousness at the macro-level is perfectly correlated with maximal integrated information, as the integrated information theory proposes. The philosophical solutions to the mind-body problem outlined above will account for this putative scientific fact in terms of the following hypotheses:

1. *Materialism* – Consciousness is identical with maximal integrated information,
2. *Panpsychism* – Conscious particles combine when there's maximal integrated information.

3. *Dualism* – The psycho-physical laws (i.e. the fundamental causal principles connecting physical reality to non-physical consciousness) ensure that consciousness arises when there is maximal integrated information.

As far as observation and introspection are concerned, these three hypotheses give the same result: macro-level consciousness is correlated with maximal integrated information. If we can't decide between these hypotheses with an experiment, how should we decide between them? At first glance, you might think materialism is the more parsimonious option, eschewing the need for either conscious particles or psycho-physical laws. However, whilst materialists aspire to reduce consciousness to physical processes, panpsychists aspire to account for physical reality itself in terms of underlying facts about consciousness (more on this below). Each, if successful, bottoms out at a single primitive – *physical reality* for the materialist, *consciousness* for the panpsychist – and hence they are equally parsimonious.

Just based on considerations of parsimony, then, panpsychism and materialism are on a par, and either is to be preferred over the dualist position that involves two primitives – consciousness *and* physical reality – rather than one. If there were no other considerations to appeal to, perhaps a position of agnosticism between materialism and panpsychism would be called for. But if, like me, you think the knowledge and conceivability arguments successfully demonstrate the incoherence of the materialist attempt to reduce consciousness to physical processes, then that leaves you with panpsychism.

This, in very broad brushstrokes (see *Consciousness and Fundamental Reality* for the details), is my case for panpsychism. When Barwich asks on p. 147 'So, how can you prove or disprove the metaphysical tenets of panpsychism?', this would have been a good moment to outline my case. Instead, she merely says that '[t]he point is: you cannot—moreover, you need not. Panpsychism is offered as a descriptive, not a productive worldview (meaning, it describes appearances under the lens of common intuition, although without further investigation or by offering any constructive approaches to better understanding of the phenomena it describes).'

This totally ignores the explanatory benefits contemporary panpsychists argue that their position has, namely that it provides a way of accounting for physical reality as emergent from underlying facts about consciousness. This builds on the insights of Bertrand Russell in his 1927 book *The Analysis of Matter*. In this book, Russell points out that the mathematical nature of physics entails that, in a certain sense, physics doesn't tell us the nature of fundamental reality: it merely describes its mathematical structure. As far as physics is concerned, fundamental reality could turn out to be *anything*, so long as it has the right mathematical structure. The Russellian panpsychist exploits this by theorising that fundamental reality consists of networks of very simple conscious entities, interacting in simple predictable ways. Through their interactions, they realise certain patterns and mathematical structures. The claim of the Russellian panpsychist is that those mathematical structures *just are* what we call physics, and hence physical reality emerges from these basic facts about consciousness. Whereas the materialist has failed to explain how consciousness could emerge from matter, the Russellian panpsychist has provided an explanation of how matter could emerge from consciousness.

Barwich says absolutely nothing about this crucial core of the panpsychist explanatory project. My suspicion, based on the fact that the references to my work don't go past page 11 of my popular-level book, is that she didn't get to chapter 4 where this is covered (never mind reading any of my academic work). This huge gap in Barwich's understanding of the panpsychist position becomes important when she says the following:

Shouldn't that theory entail and build on a sufficient understanding of electrons, their properties, and their past and present scientific investigation? Physicists (Hossenfelder 2020) publicly criticized the absence and conflict with modern physical theories of fundamental particles in panpsychism (p. 152).

Russellian panpsychists *do* build on an understanding of the scientific characterisation of electrons, in particular that physics limits itself to describing them in purely mathematical terms. The physicist Max Tegmark (2014) infers from this there is nothing more to physical reality than the mathematical nature physics ascribes to it. Panpsychists, in contrast, hold that that the mathematical structures of physical reality are realised by underlying facts about the interactions of conscious entities, that it is consciousness that 'breathes fire into the equations,' to use Stephen Hawking's memorable term.²

Moreover, as I've discussed elsewhere, Hossenfelder fails to hit target simply because she doesn't know the details of the contemporary panpsychist position, which she says she learnt about from 'pamphlets.'³ She assumes that the panpsychist is postulating consciousness *in addition* to the physical properties of electrons, such as mass, spin and charge. In fact, the panpsychist leaves the standard model of particle physics untouched, postulating a layer *underneath* the mathematical structures of physics. Hossenfelder's objection is akin to a biologist who doubts the existence of the quantum wave function because it doesn't feature in cellular biology. This would not be a good argument, as the quantum wave function exists at a more fundamental level than cellular biology, and hence is bound not to crop up in cellular biology. Likewise, the fundamental consciousness postulating by the panpsychist exists, if it exists, beneath the mathematical structures of physics.

I suspect that if she'd read up to chapter 4, Barwich would object that this extra layer of conscious reality is doing no 'productive' work. Well, it's not doing any extra *scientific* work. But my view is that the mind-body problem (as distinct from the science of consciousness) is a philosophical rather than a scientific challenge, and Barwich hasn't given us any argument to the contrary. Russellian panpsychism provides an explanatory reduction of physical reality in terms of underlying facts about consciousness, which results in a general ontological position as parsimonious as materialism. If materialism is then ruled out, e.g., by the knowledge argument, then panpsychism looks to be the preferred philosophical solution to the mind-body problem (of course, that's a big 'if,' but Barwich has not even considered the arguments in support of that antecedent).

Barwich refers to many examples of how neuroscience has deepened our understanding of consciousness, suggesting that this shows that I am 'remarkably wrong' in holding that neuroscience has not shed any light on how the brain produces consciousness (p. 47). But it is a caricature of my position to characterise it as the view that science has nothing to offer in understanding consciousness. As I state above, and as is explicit in *Galileo's Error* (p. 35), my view is that neuroscience can shed light on *which* kinds of brain activity go along with which kinds of conscious experience, but cannot explain *why* those specific kinds of brain activity go along with those kinds (or indeed any kinds) of conscious experience. Indeed, one of the examples Barwich refers to, namely, Perry and Gazzaniga's split brain research, is one I discuss in *Galileo's Error*. At the very least, Barwich ought to have argued that the examples she refers to are not merely cases of what I explicitly say neuroscience *can* do, but rather cases of what I say neuroscience *cannot* do.

² Hawking 1998.

³ This was a 2019 blog post called 'Electrons Don't Think': <https://backreaction.blogspot.com/2019/01/electrons-dont-think.html> I respond in Goff 2023.

In fact, I don't know of any anti-materialists in academic philosophy who have been anything but highly enthusiastic about developments in the science of consciousness, and it therefore seems to me inaccurate for Barwich to say, without a single supporting citation, that 'Philosophers initially responded with hostility toward this new research program, arguing that certain things lie outside the grasp of experimental science—with consciousness at the front and center' or that in my work 'the material reality of science seems dismissed right from the outset (regardless of whether it is physics, biology, psychology, or neuroscience)' (p. 152). I am not dismissing scientific data, it's just not relevant to the question of whether materialism or panpsychism is true given that these views are empirically equivalent.

There are many ad hominem in this article, e.g., 'panpsychism embodies a metaphysical attitude that, in (too) many of its current proposals, appears to tell us more about how a number of people want to think and what they want to believe about consciousness than clarify or explain something about what consciousness is.' (p. 151). Anyone can speculate about the hidden motives behind the views of a given philosopher, but it has no impact on the arguments as they are presented. Barwich also often refers to panpsychism as a 'personal belief,' seeming to contrast this with something that it's empirically supported (p. 146, 151). This is false dichotomy. Galileo used a philosophical thought experiment, rather than an empirical experiment, to rule out the Aristotelian hypothesis that heavier objects fall faster than lighter ones, but it would not be appropriate to describe this as merely his 'personal belief,' as though it lacked objective rational support. Similarly, in contrast to my rejection of marmite, my rejection of materialism is not based on personal taste but on my conviction that certain thought experiments demonstrate its incoherence. I might be right, I might be wrong, but either way my view is based on rational arguments not 'personal belief.'

Having said all of the above, there is much to like in Barwich's paper. In particular, the discussions of the nature of science and the importance of science literacy are insightful and timely. I'm actually a huge fan of Barwich's work. I just don't think these discussions have much to do with my philosophical case for panpsychism. She tries to link the two in this passage:

The point here is to highlight that scientific thinking cannot be analyzed somewhat ahistorically as an unchanging cognitive procedure from Galileo to today that's birthed by the scientific revolution without affecting human cognition and the mental mechanisms with which we engage with the world (as it is, for instance, propagated in Goff 2019). The activity of science and the participation by cognitive agents within its practices consists of much more than culturally inherited information that gets passively taught and rethought. Essentially, science is demarcated by the active development of an evolving cognitive skillset (p. 159).

At no point do I claim that science has not changed in its approach for the last 500 years. My claim is merely that what has remained constant since Galileo is the purely quantitative conception of the nature of matter, and Barwich says nothing to dispute that. Questioning my claim in chapter 1 that science is often as much to do with 'dreaming up possibilities' as empirical data, Barwich goes on to say:

...current science does not fail in "dreaming up possibilities," as science isn't about imagining answers. Science is about cultivating curiosity about how the world works beyond preconceived philosophical intuitions and situated beliefs, which it does by developing more sophisticated and sometimes new questions. Georg Cantor revolutionized mathematical thinking by asking whether there are different infinities. Albert Einstein redefined our thinking about time and space by reimagining the question of what simultaneity is. Charles

Darwin shifted biological perspectives about development and inheritance by introducing the idea of probability into the organic world (p. 159).

Well, Darwin did 'dream up' the possibility that organisms might have come about through natural selection, thus identifying a new option in conceptual space between intelligent design and the hypothesis that the complex functioning of organism arose by chance. Likewise, Russell 'dreamed up' the possibility of grounding the mathematical structures of physics in underlying facts about mind or consciousness, thus identifying a new option in conceptual space between dualism and materialism. There are many differences between these two examples, but they both do involve 'dreaming up possibilities.'

More generally, it seems to me somewhat contrived to insist on framing the insights of Cantor and Einstein in terms of 'questions' rather than 'answers.' Cantor didn't merely *ask* whether there are different sizes of infinity, he managed to *prove* that there are. I guess you could frame Einstein's innovation in special relativity as all about a 'question' if you like, but essentially Einstein dared to imagine the possibility that what is simultaneous to what is relative to a frame of reference rather than absolute (because he dared to imagine the possibility that a relativity principle applied to electromagnetism). You could equally frame my view as 'reimagining the question' of whether the physical or the experiential is fundamental.

I very much like Barwich's conception of science as 'an evolving cognitive skillset,' and her view that we must 'adapt our methods to the puzzle, not the other way around' (p. 160, taken from Barwich 2020, 311). If we allow for a significant part of 'science' being a priori rather than directly empirical, then I would say these descriptions capture what I'm doing. I have the impression that Barwich adopts the scientific view that the a priori has a minimal role if any in the project of finding out the nature of reality. But, again, we don't find in this paper an *argument* for that view.

We do find one argument in the paper against one of my views, specifically my view that we know something of the nature of experiences through introspection. This thesis has come to have the perhaps unfortunate name of 'Revelation' (Stoljar 2009: 115). Barwich questions this thesis, saying 'just because we experience consciousness directly, do we necessarily know its nature?', before providing examples of empirical progress showing that memory works differently to how we thought and that flavours we experience as located in the mouth in fact have an olfactory cause (p. 150).

It's not clear to me in either case that the phenomenon in question is directly experienced. Many think we're directly aware of flavour experiences but few think we're directly aware of their physiological causes. Moreover, I have nowhere argued that Revelation must be true 'just because we experience consciousness directly.' Rather I have an extensive argument for Revelation in chapter 5 of *Consciousness and Fundamental Reality* (also in Goff 2011, 2015), an argument which Barwich doesn't consider. Again, the objection misses the mark because it proceeds without any real knowledge of my work (beyond chapter 1 of my popular-level book, a review of that book in the popular press, and a popular-level exchange of letters with a scientist).⁴

I have the feeling Barwich is just not interested in these metaphysical questions as to whether panpsychism or materialism or dualism is correct. That's fine; different people are interested in different things, and Barwich is doing incredible work in a different area of philosophy. But I think if

⁴ Barwich also reference the 'Panpsychism' entry in the Stanford Encyclopedia of Philosophy (Goff et al), however, there is no consideration of the arguments for panpsychism contained there.

you're going to disagree with a particular philosopher in print, then you should engage with the arguments they give in support of their position.

References

- Barwich, Ann-Sophie (2020) *Smellosophy: What the Nose Tells the Mind*, Harvard University Press.
- Barwich, Ann-Sophie (2022) 'Between electrical light switches and panpsychism: Scientism and the responsibilities of the humanities in the twenty-first century,' in Moti Mizrahi (Ed.) *For and Against Scientism Science, Methodology, and the Future of Philosophy*, Rowman & Littlefield Publishers
- Goff, Philip (2011) 'A posteriori physicalists get our phenomenal concepts wrong,' *Australasian Journal of Philosophy*, 89: 2, 191-209.
- Goff, Philip (2015) 'Real acquaintance and physicalism,' in P. Coates & S. Coleman (Eds.) *Phenomenal Qualities: Sense, Perception, and Consciousness*, Coates, Oxford University Press.
- Goff, Philip (2017) *Consciousness and Fundamental Reality*, Oxford University Press.
- Goff, Philip (2019) *Galileo's Error: Foundations for a New Science of Consciousness*, Pantheon.
- Goff, Philip (2023) *Why? The Purpose of the Universe*, Oxford University Press.
- Goff, Philip, William Seager, & Sean Allen-Hermanson (2017) 'Panpsychism,' in Ed. Zalta (Ed.) *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/panpsychism/>.
- Harris, Sam (2012) *The Moral Landscape: How Science can Determine Human Values*, Bantam Press.
- Hawking, Stephen (1988) *A Brief History of Time: From the Big Bang to Black Holes*, New York: Bantam.
- Tegmark, Max (2014) *Our Mathematical Universe: My Quest for the Ultimate Nature of Reality*, Knopf.
- Russell, Bertrand (1927) *The Analysis of Matter*, Kegan Paul.
- Stoljar, Daniel (2009) 'The Argument from Revelation,' in David Braddon-Mitchell and Robert Nola (Eds.) *Conceptual Analysis and Philosophical Naturalism*, MIT Press, 112–38.