

ILLUSIONISM IS NO TRICK

KEITH FRANKISH

Introduction

Paul Stenner is not persuaded by the illusionist view presented in my interview with Katarína Sklutová. In fact, he thinks I am a trickster, who is not only defending a position I call ‘illusionism’ but perpetrating an intellectual illusion myself. He suggests that my arguments against qualia are a piece of misdirection, designed to deceive readers into accepting that we should eliminate the notions of mind and consciousness altogether and think of ourselves (if only we could think!) as unfeeling machines. Stenner thinks I am motivated in this aim by a desire to defend a crude form of materialism, which ignores modern physics.

I’m puzzled. I don’t recognize Stenner’s account of what I am doing or why I am doing it. Indeed, after reading Stenner’s reply, I wasn’t sure whether he disagrees with my actual views, as opposed to the caricature he presents. Rather than speculate about this, I shall use this reply to clarify my position and the background to it. Perhaps I didn’t express myself carefully enough in the interview. If so, let me try to do better.¹

I shall begin by making some preliminary points about the background to my view of consciousness and then sketch both the view I endorse and the one I reject. Perhaps this will pinpoint the substantive issues on which Stenner and I disagree and help readers decide where their own sympathies lie.

Preliminaries

The first thing to say is that I do not claim that minds and mental states are illusory. I do not subscribe to eliminativism about folk psychology of the sort defended by, for example, Paul Churchland (Churchland, 1981). In fact, most of my earlier research was devoted to defending realism about belief/desire psychology against the eliminativist challenge. (For

¹ The interview was conducted by email, which explains my ability to use parentheses in my replies — a fact remarked on by Stenner.

the full story, see my 2004.) So there is not, as Stenner suggests, an inconsistency in my claiming to *believe* my own theory. My eliminativism is restricted to certain supposed mental properties — the ones philosophers call *qualia* or *phenomenal properties* — and the distinctive form of consciousness they are supposed to constitute. Illusionism is precisely the view that qualia are an introspective illusion.

Second, I am not denying that we have conscious experiences and a subjective take on the world. There is a big difference between conscious experience and subliminal perception. Nor am I claiming that people do not have a subjective take on the world, consisting in a complex of sensitivities and a plethora of covert and overt responses. Such takes are very real, and a person's sincere reports of them — reports of how things *seem* to them — are reliable and highly informative expressions of their psychological state. I am not suggesting that people are wrong about what they experience. Nor, finally, am I denying that conscious experiences have the moral significance we take them to have. I do not think that people are unfeeling machines. (Though I think that suitably complex machines could feel.)²

The issue is not whether people are conscious, but what it *means* to be conscious. And here I do think that people can be wrong. Being conscious doesn't automatically confer a correct theory of what consciousness is, any more than being alive confers a correct theory of what life is. A person may know they are in the state we call 'pain' without having a full understanding of what pain is. People may differ, for example, on *where* pain is located — whether it is in their body or their mind.

Third, materialism. Stenner claims that I am committed both to an outdated conception of the physical (similar to Galileo's!) and to a strong form of reductionism, of the sort associated with the logical positivists. I didn't say much about this in the interview (Stenner supports his interpretation by quoting a few words out of context), so I shall be more explicit here.

I would describe myself as a physicalist in the following sense. I take it that physics offers our best account of the fundamental entities — processes, fields, forces, etc. — from which everything is constituted. This is just taking physics seriously, since providing such an account is precisely what physics aims to do. Being a physicalist in this sense involves holding that everything is physical in the broad sense of being ultimately grounded in the entities posited by physics.³ But it does not involve believing that everything is material in an everyday sense. Modern physics paints a wildly counterintuitive picture of fundamental reality. (According to one major school of thought, fundamental reality can be described as a universal wave function unfolding in accordance with the Schrödinger equation.) And as physics develops, physicalism will develop with it. I do, however, assume — as a working hypothesis — that physics will not need to posit fundamental *mental* properties and forces. It certainly shows no sign of doing so.⁴

² It is true that some interviews with me have appeared under titles indicating that I think consciousness is an illusion, but these titles were not chosen by me. I always make it clear that the only form of consciousness I deny is the putative *phenomenal* form that consists in possession of qualia.

³ More accurately, every *spatiotemporally located* thing; abstract entities are another matter.

⁴ It is sometimes said that physicists hold that the mental act of observation plays an essential role in collapsing the wave function, but this is a misapprehension. Any kind of measurement can trigger collapse, including one made by a simple mechanical device. And there is a strong case for thinking that collapse does not really occur anyway (e.g., Carroll, 2019).

Physicalism in this sense is an ontological claim — a claim about the nature of reality — not an epistemic claim about how we can know that reality. I am not claiming that physics can *explain* everything. Of course, it cannot explain higher-level events — geological, historical, sociological, psychological, and so on.⁵ It doesn't have the right concepts and theoretical frameworks. Moreover, I reject the strong reductionism advocated by Ernst Nagel and others, according to which higher-level theories can be logically derived from basic physical principles with the aid of bridge laws (Nagel, 1961). This view is now widely rejected — not least because different instances of the same type of higher-level process can be grounded in different physical processes. Physics cannot replace the special sciences, even if the things the special sciences talk about are ultimately grounded in the things physics talks about. (For useful discussion of these issues, see Tahko, 2021.)

Though I reject strong reductionism, I do think that we can hope to find reductive explanations of specific instances of higher-level features, including psychological ones. That is, we can hope to identify the mechanisms in virtue of which the features exist. However, such explanations will rarely be neat. They will often be *partial*, appealing to mechanisms that are themselves highly complex; *local*, applying only to certain instances, such as those in a specific group of organisms; *disunified*, citing many independent mechanisms; and *complex*, involving multiple factors at various scales. What I reject is the idea that there is an *in principle* explanatory gap between neurobiology and aspects of psychology, as some have argued (e.g., Levine, 1983). To assume so — at least in our present state of knowledge — would be to evade a central explanatory task for cognitive science.

That's my default attitude, but I am not dogmatically committed to it, and I would modify it if there were a strong enough case for doing so. I made this claim in the interview, but Stenner thinks this is a trick too, saying that I wouldn't accept an argument as strong enough unless it was a physicalist one (Stenner, 2022, [ms, p.4]). I don't understand this. I am saying that I would give up physicalism if there were a strong enough *anti*-physicalist case, and I mean it. Here's an example of such a case (not the only possible one, of course). Suppose scientists were to find solid experimental evidence that certain psychological events — conscious decisions, say — produce changes in the physical state of the brain that cannot be explained in terms of known physical forces. This would be good evidence for the existence of irreducible mental forces, and thus against the sort of physicalism I described.

What I do *not* accept as evidence against physicalism are intuitions, based on introspection and armchair reflection. Without data and a theoretical framework, intuitions count for little in science, even if some philosophers take them seriously.

Two views of consciousness

With these preliminaries out of the way, I shall now sketch two views of what it is to be conscious, the one I endorse and the one I reject. I shall paint with a broad brush in order to bring out the contrast.

⁵ Talk of *levels* shouldn't be taken literally, and it would be more accurate to speak of larger-scale patterns. But the term is widely used and is convenient.

The view I endorse is what I shall call the *reactive view*. It is this. Consciousness involves reacting in a certain way to features of the world or one's own body — the features we call colours, smells, sounds, pains, and so on. These features are physical ones (surface properties, airborne compounds, pressure waves, tissue damage, and so on), though this is not obvious to us. For you to be conscious of such features is for you to be actively sensing their presence and for them to be affecting you in a wide-ranging way — evoking a myriad covert and overt responses, physiological, psychological, and behavioural. The cumulative effect of these reactions is to make you directly aware of the features and enable you to respond to them as a person — to think about them, act upon them, tell others about them, and so on.

When I say 'you' here, I mean the whole organism, the person, constituted by all the biological and psychological systems in your body; no dualism is supposed. And when I say that being conscious of things involves you being *directly aware* of them, I do not mean that there are no mechanisms involved. Features affect you via proximal stimuli that impinge on your sense organs and initiate hugely complex waves of neural processes. But you are not aware of these mechanisms; they operate at a *subpersonal* level. The mechanisms lock you on to distal features of the world in a tight dynamic of sensitivity and response, whose effect is to put *you* — the whole organism — into a direct personal relation to the features of the kind we call *experiencing*. Of course, these mechanisms can go wrong or be tricked, as happens when we dream, hallucinate, or misperceive things. In these cases, we enter reactive states just like those we enter when we react to real features, and we call these states conscious experiences too.

On this view, then, conscious experience is constituted by a complex of activated sensitivities and reactions. These sensitivities and reactions are *subjective* — they are shaped by the subject's biology, psychology, and personal history — but they are not essentially private, and we can hope to provide reductive explanations of them (of the limited kind described). *Subjective* doesn't entail *private*. My views about art are subjective, but they needn't be private.

The view I reject I shall call the *intrinsic view*. According to this, the reactive view misses out something essential. Consciousness isn't just a complex reaction to features of the world; there is an *intrinsic* nature to conscious experience that cannot be accounted for by world-directed processes of sensitivity and response. We could be sensitive to surface features of the kind we call 'red', and respond to them in all the ways we do, without having a conscious experience of red at all. The experienced quality of redness is something extra — a mental feature, which accompanies the sensory processes. The same goes for all other experiences. In each case, the sensitivities and reactions are not sufficient for conscious experience. We can imagine a *zombie* which responds to the world in exactly the same way but experiences nothing. These experienced qualities are what philosophers call 'qualia' or 'phenomenal properties', and accounting for them is the notorious 'hard problem' of consciousness. The classic modern statement of this view is David Chalmers's hugely influential book *The Conscious Mind* (Chalmers, 1996).

Qualia and illusion

The choice between the two views depends on whether you think qualia exist — whether you are a *qualia realist* — so I shall say a little about this. Stenner doubts that anyone is a qualia realist (‘What is a “mental quality” and who believes in them anyway?’; Stenner, 2022, [ms, p. 5]).⁶ If only it were so! Belief in qualia is rife among philosophers of mind — and not uncommon among neuroscientists. (For an introduction to the topic, see Tye, 2021.) Indeed, many take it to be the only way to think of consciousness:

We can say that a being is conscious if there is *something it is like* to be that being, to use a phrase made famous by Thomas Nagel. Similarly, a mental state is conscious if there is something it is like to be in that mental state. To put it another way, we can say that a mental state is conscious if it has a *qualitative feel* — an associated quality of experience. These qualitative feels are also known as phenomenal qualities, or *qualia* for short. The problem of explaining these phenomenal qualities is just the problem of explaining consciousness. This is the really hard part of the mind–body problem. (Chalmers, 1996, p. 4)

That is, for a mental state to be like something is for it to have a distinctive property — a what-it-is-likeness, a quale (singular of qualia).

Contemporary philosophy of consciousness centres obsessively on qualia and their ontological status. Many philosophers argue that qualia are not physical properties (in the broad sense of being grounded in the entities of physics) and conclude that we must expand our inventory of fundamental entities to accommodate them (e.g., Chalmers, 1996). More recently, there has been a surge in popularity for the *panpsychist* view that qualia are present in all physical entities (e.g., Goff, 2019).

It is in this context that illusionism must be seen. Illusionists reject the intrinsic view and argue that we can explain our intuitions about what-it-is-likeness within a reactive view. The general idea (which can be fleshed out in various ways) is this. As well as being sensitive to features of the world, we are also sensitive to our own processes of sensitivity and the reactions they evoke — thanks, presumably, to internal monitoring and modelling systems within the brain. We are able to recognize different patterns of sensitivity and reaction — different experiences — when they occur in us. But our introspective abilities are limited; we have no access to the detail of the patterns and can only describe them in simple, schematic terms — by their modality, their relation to each other, and their valence, for example. We gesture at this by talking of what our experiences are *like*, and we are tempted — especially if we are philosophers — to think that in doing so we are referring to some special property they have, their what-it-is-likeness. That is, we are tempted to mistake indirect, partial awareness of complex processes for direct, complete awareness of simple properties. This is where the illusion lies. We misinterpret the evidence of introspection in much the same way that an audience at a magic show misinterprets the evidence of their eyes.

⁶ Stenner derides my — charitable — suggestion that belief in qualia is a ‘natural’ inference to draw from reflection on cases where we seem to perceive qualities that aren’t really there, as when we dream or hallucinate. Whether or not the inference is natural, I agree with Stenner that it is a bad one. In such cases we are not really perceiving anything, but simply entering similar sensory/reactive states to those we enter when we do perceive things.

Motivation

How do we decide between the two pictures? We can't do it simply by attending to our own experiences — introspecting. For the two views offer rival accounts of what introspection does and how we should interpret its deliverances, one saying that it affords opaque access to complex neural processes, the other that it transparently reveals simple mental qualities. Consulting introspection itself won't settle this. We need to look at wider theoretical considerations.⁷

I shall not go into the arguments here (for some of them, see Dennett, 1988, 1991, 2005; Frankish, 2016a, 2016b). Instead, I shall conclude by saying something about the motivation for illusionism. Are illusionists motivated simply by a commitment to physicalism?

There are certainly grounds for thinking that qualia realism is incompatible with physicalism. Anti-physicalist realists such as Chalmers have some good arguments. But the reason for preferring illusionism to anti-physicalist realism isn't a dogmatic faith in physicalism. It is an awareness of the theoretical costs of treating consciousness as nonphysical. In particular, it is hard to see how nonphysical qualia could have any effects in the physical world. Unless there are fundamental forces currently unknown to physics, it seems that consciousness must be causally idle — a conclusion some qualia realists have accepted (e.g., Jackson, 1982). This is at least *some* motivation for considering illusionism.

Moreover, there are reasons for rejecting qualia realism that have nothing to do with physicalism. As Daniel Dennett has shown, there are many internal problems with the intrinsic picture, relating, for example, to our knowledge of our own qualia (Dennett, 1988, 1992, 2005). Such knowledge would be partially dependent on memory, and there will be situations where it would be impossible in principle for a subject to say what their qualia were. Since qualia are not publicly detectable, this would mean that there are facts about your consciousness that not even you could know. Again, this is some reason for exploring alternative views.

Finally, qualia realism excludes consciousness from scientific investigation. Qualia are wholly private, and we have no objective way of describing them, comparing them, or testing claims about them. Perhaps there are things that lie outside the scope of scientific investigation, but we shouldn't be hasty to locate things there.

Conclusion

Stenner's hostility to illusionism seems to stem from the impression that it is a radical form of eliminativism, motivated solely by a crude and strongly reductionist form of materialism. I have tried to set the record straight, showing that illusionism is a modest, ontologically deflationary view, which brings consciousness within the scope of scientific investigation. Now that I have clarified the scope of the thesis and the motives for considering it, perhaps Stenner will take another look. Illusionism is no trick, and it needs no trickery to make it attractive.

⁷ It might be said that introspection is self-validating — that it reveals itself as transparently revealing mental reality. But that claim is just another piece of theorizing.

References

- Carroll, S. (2019). *Something deeply hidden: Quantum worlds and the emergence of spacetime*. Dutton.
- Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford University Press.
- Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy*, 78(2), 67–90.
- Dennett, D. C. (1988). Quining qualia. In A. J. Marcel and E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 42–77). Oxford University Press.
- Dennett, D. C. (1991). *Consciousness explained*. Little, Brown, and Co.
- Dennett, D. C. (2005). *Sweet dreams: Philosophical obstacles to a science of consciousness*. MIT Press.
- Frankish, K. (2004). *Mind and supermind*. Cambridge University Press.
- Frankish, K. (2016a). Illusionism as a theory of consciousness. *Journal of Consciousness Studies*, 23(11–12), 11–39.
- Frankish, K. (2016b). Not disillusioned: Reply to commentators. *Journal of Consciousness Studies*, 23(11–12), 256–289.
- Goff, P. (2019). *Galileo's error: Foundations for a new science of consciousness*. Pantheon Books.
- Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly*, 32(127), 127–136.
- Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64(4), 354–361.
- Nagel, E. (1961). *The structure of science: Problems in the logic of scientific explanation*. Harcourt, Brace & World.
- Stenner, P. (2022). 'Trust me, I'm an illusionist.' A critical response to Keith Frankish's Illusionism and its place in contemporary philosophy of mind. *Human Affairs*, 32(3), 311–320.
- Tahko, T. E. (2021). *Unity of science*. Cambridge University Press.
- Tye, M. (2021). Qualia. In E. N. Zalta (Ed.) *The Stanford encyclopedia of philosophy* (Fall 2021 edition). <https://plato.stanford.edu/archives/fall2021/entries/qualia/>.

Department of Philosophy

45 Victoria Street

Sheffield

S3 7QB

United Kingdom

Email: k.frankish@sheffield.ac.uk

ORCID: 0000-0003-0968-2944