



The consciousness illusion

Phenomenal consciousness is a fiction written by our brains to help us track the impact that the world makes on us

Keith Frankish

In the movie *The Matrix* (1999), Morpheus offers Neo a red pill. If he takes it, he will discover that reality as he knows it is an illusion created by machine overlords to keep humans enslaved. I am going to offer you a different pill, which – if it works – will convince you that your own consciousness is a sort of illusion, a fiction created by your brain to help you keep track of its activities. This view – which I call illusionism – is widely considered absurd (it's been described by Galen Strawson as 'the silliest

claim ever made'), but it has able defenders (pre-eminently Daniel Dennett), and I want to persuade you that it isn't absurd and might well be true. Are you ready to see how deep the rabbit hole goes?

The first task is to be clear what we're talking about. The term 'consciousness' is used in different ways, and when I claim that consciousness is illusory, I mean it only in one specific sense. We can home in on our target with an example. I'll take vision, but any other sense would do as well. Suppose you have good sight and are focusing on a red apple directly in front of you in good lighting. You are now in a certain mental state, which we can call having a conscious visual experience of the apple. You wouldn't be in this state if you were unconscious or asleep (though if you were dreaming, you might be in a similar state), or if you had not noticed the apple, or had noticed it only in a fleeting, subliminal way. Our lives are filled with such experiences, and no one suggests that they are not real. The question is what is involved in having such experiences, and whether it involves consciousness in a more specific sense.

So, what is involved in consciously experiencing the apple? Well, lots of things. You are acquiring a mass of information about the apple – fine-grained details of its shape, colour, texture, location, distance and so on. You are recognising what kind of thing it is (a solid object, a piece of fruit, an apple, a Red Delicious) and forming corresponding beliefs (that there is a thing of this kind in front of you). You are recognising ways you might interact with the apple and the opportunities or threats it offers – what psychologists call its affordances. You recognise the apple as something you might pick up, juggle with, eat, cook and so on. You are also getting ready to react. You are forming expectations about the apple (that it won't move or attack) and inclinations to respond to it (you might feel an urge to grab it and take a bite). Memories and associations are being evoked, perhaps affecting your mood or setting your thoughts on a different track. You don't explicitly think about all these things, of course, but you would report many of them if questioned, and we know from experimental work that a vast array of sensitivities and associations are triggered during a conscious experience, priming us to react to future stimuli and collectively determining the experience's significance for us.

Neuroscientists are beginning to understand the brain processes underlying all this. To put it simply, light reflected from the apple stimulates light-sensitive cells in the retina, sending trains of electrochemical impulses along the optic nerve to the lateral geniculate nucleus and then on to the visual cortex at the back of the brain. Here these signals trigger activity in hierarchically organised groups of cells specialised for the detection of increasingly complex features (edges, colours, motion, faces and so on). When you attend to what you are seeing, this visual information is 'globally

broadcast' to mental systems involved in memory, reasoning, emotion and decision making, generating the host of effects mentioned. This process of global broadcast is called access consciousness, since it makes sensory information accessible to the rest of the mind, and thus to 'you' – the person constituted by these embodied mental systems. Again, I don't deny the reality of consciousness in this sense.

As yet, we have only a sketchy understanding of access consciousness, and there are many controversies over the details of the neural systems involved, but in time we should be able to fill in the picture and settle the disputes. Yet, many philosophers would say, even then we wouldn't have a full understanding of consciousness. For, they would say, something else is happening, alongside all this information processing. It is like something for you to see the apple. You are having visual sensations of its redness, brightness, shininess and so on, each with a distinctive quality to it. These seem to be further aspects of the experience, in addition to all the sensitivities and reactions, and we can imagine (can't we?) all the latter occurring without it actually being like anything for you to see the apple. It could have been 'all ... dark inside' in David Chalmers's phrase from his book <https://global.oup.com/ushe/product/the-conscious-mind-9780195117899?cc=gb&lang=en&> *The Conscious Mind* (1996). (Perhaps that's how it would be for robots?)

The point is not restricted to vision, of course. We also have conscious experiences involving our other senses. Think of attentively listening to a flute, smelling a rose, tasting a wine or feeling a velvety fabric. Think, too, of attending to bodily experiences, such as a stubbed toe, a migraine headache or a surge of euphoria. In each case, we can tell a similar story to the one about vision. Again, the experience involves acquiring information (about sounds, substances, textures or the state of one's body), identifying objects, forming beliefs, sensing opportunities for action, generating expectations and reactive dispositions, and the triggering of memories, associations and emotions. Again, this involves access consciousness. Sensory signals are routed to specialist processing areas of the brain, and the information extracted is then globally broadcast to other mental systems that produce the various effects. And, again, there seems to be more to experience than all these sensitivities and reactions. It is like something to hear the flute, smell the rose, taste the wine and so on. Each experience seems to have a distinctive sensory quality to it. The experience of a stubbed toe, for example, carries information about bodily damage and will trigger a host of negative associations and reactions, but (the story goes) it also has a qualitative aspect – a pure awfulness.

Here you might want to object (rightly, I think) that these sensory qualities don't seem to be features of our experiences but of the things we experience. When you attend to the apple, the red quality you experience seems to be a feature of the apple, which causes the reactions in you. (You believe the apple is red because it looks red.) Similarly, the sound seems to be in the air, the taste in the wine, the pain in your toe, and so on. But it is generally agreed that this can't be right. For science tells us that objects don't have such qualitative properties, just complex physical ones of the sort described by physics and chemistry. The atoms that make up the skin of the apple aren't red. Considered as properties of external things, colours are light-reflective surface features, sounds are vibrations in the air, tastes and smells are chemical compounds, and the pain in your toe is cell damage. It seems, then, that the qualities of colour, sound, pain and so on exist only in our minds, as properties of our experiences. Philosophers refer to these subjective qualities of experience as 'qualia' or 'phenomenal properties', and they say that creatures whose experiences have them are phenomenally conscious.

If our sensations are nonphysical features, then they wouldn't have any effects, even on our own thoughts

It is phenomenal consciousness that I believe is illusory. For science finds nothing qualitative in our brains, any more than in the world outside. The atoms in your brain aren't coloured and they don't compose a colourful inner image. (And even if they did, there is no inner eye to see it.) Nor do they have any other qualitative properties. There are no inner sounds, smells, tastes and pains, and no inner observer to experience them if there were. It is true that cognitive scientists talk of there being representations in the brain. But by this they don't mean inner pictures or copies that we observe instead of observing the world directly. They mean patterns of neuron-firing that respond to specific features of the world and that the brain uses to construct models of its environment. Representations in this sense are not things we are aware of; rather, they are parts of the machinery that makes us aware of things. By modelling the world, your brain creates sensitivities and dispositions that put 'you' – the person – in direct contact with the world. Such representations need not share properties with the things they represent. The mental representation of redness need not be red, any more than the word 'red' or the numeral that denotes red in a painting-by-numbers kit.

If you can't shake your conviction that phenomenal consciousness is real, then there are two broad options here. One is to say that phenomenal consciousness is an extra feature of the brain, in addition to the physical properties described by science. (This

view is called property dualism, since it holds that the brain has dual properties, physical and nonphysical.) A big problem for this view is that it threatens to make phenomenal consciousness ineffective. There is absolutely no reliable evidence for nonphysical effects in the physical world – no confirmed cases where a nonphysical feature diverts an electron, triggers a chemical reaction, makes a neuron fire, or produces any other physical change. (Indeed, such changes would violate the basic principle of the conservation of energy.) If our sensations are nonphysical features, then it looks like they don't have any effects on anything, even on our own thoughts and reactions. We would think and act in exactly the same way (including believing that we are phenomenally conscious) if we didn't have them or had completely different ones. This is a strange conclusion.

The other realist option is to say that, despite appearances, phenomenal consciousness is a purely physical process. Phenomenal states, such as the sensation of red, the smell of a rose or the pain of stubbed toe are nothing more than brain states, which could in principle be observed by other people. A brain surgeon could literally see your pains! (This is a form of physicalism.) Defenders of this view argue that the reason our sensations don't seem to be brain states is that we have a special perspective on them. We can recognise them 'from the inside', via introspection. By this, they don't mean an inner eye but an internal monitoring system that registers information about our brain states. When you attend to the apple, your introspective mechanisms detect the resulting brain state and make information about it available to other mental systems, thereby making 'you' aware of it. But (the argument goes) introspection gives you very limited information about the state. It doesn't represent it *as* a brain state but just as something occurring in you, which you can describe only as the 'feel' of seeing red. This is the orthodox physicalist view of phenomenal consciousness. It faces a problem, however. For it is not just that introspection fails to present sensations as brain states; it positively presents them as utterly unlike brain states, and this intuition doesn't go away even if one is firmly convinced that they must be brain states. It's as if introspection misrepresents our brain states to us.

This brings us to illusionism. Illusionists agree with other physicalists that our sense of having a rich phenomenal consciousness is due to introspective mechanisms. But they add that these mechanisms misrepresent their targets. Think of watching a movie. What your eyes are actually witnessing is a series of still images rapidly succeeding each other. But your visual system represents these images as a single fluid moving image. The motion is an illusion. Similarly, illusionists argue, your introspective system misrepresents complex patterns of brain activity as simple phenomenal properties. The phenomenality is an illusion.

What is happening, I believe, is something like this. The processes of access consciousness can operate without higher-level monitoring, enabling us to respond rapidly and flexibly to our environment. However, it is useful to us to have an overview or ‘edited digest’ (Dennett’s phrase) of these processes – a sense of the overall shape of our complex, dynamic interaction with the world. When we speak of what our experiences are like, we are referring to this sense, this edited digest. It works this way. Introspective mechanisms monitor access consciousness, tracking the high-level patterns of activity that encode the properties and affordances of perceived objects, and the associations, expectations, emotions and priming effects they evoke. But they don’t represent them accurately and in detail. Rather, they represent them in schematic, caricatured terms, as simple phenomenal properties, which express the general shape of the multidimensional impact that perceived objects are making on us. These representations don’t have phenomenal properties, of course. They aren’t little glowing pictures or vibrant recordings in the brain. (As I noted earlier, mental representations aren’t things we are aware of but part of the machinery that makes us aware of other things.) Rather, they are patterns of neuron-firing that signal the presence of phenomenal properties in much the same way that written words signal the presence of things. (They falsely signal these properties, of course, but in doing so they help us keep track of real physical features.) Other mental systems then use these representations for higher-level control and planning, enabling us to recognise and recall significant experiences, take steps to repeat or avoid them in future, and decide whether or not to tell others about them.

The idea, then, is that introspection tracks the impact objects make on us. The red quality you seem to experience is an expression of your response to the apple – your active ‘redding’, as the psychologist Nicholas Humphrey puts it in his book <https://www.hup.harvard.edu/catalog.php?isbn=9780674030541> *Seeing Red* (2006). However, introspection doesn’t represent phenomenal properties as properties of us but as powers in objects to create that impact. Redness is represented as a power of surfaces to affect us in a certain way, a rose smell as a power of airborne substances to affect us in another way, a stabbing pain as a power of a body part to affect us in yet another way, and so on. In each case, the character of the represented property corresponds to the nature of the impact on us. (Compare aesthetic properties, such as beauty. Our judgments of beauty reflect our reactions to things, but we think of beauty as a property of the things themselves.) This makes good design sense. The easiest way for us to remember, recall and communicate significant experiences is by picking out the objects that cause them. Our minds represent objects as decked out with illusory qualitative properties that highlight their significance for us. This perspective is not so different from the common sense one. Common sense says that qualitative properties are potent, mind-independent

features, and illusionists says that representations of qualitative properties track just such features.

We should no more expect to find phenomenal properties in our brains than to find waste baskets in our laptops

Illusionists agree, then, that there is more to conscious experience than access consciousness: there is a level of self-monitoring involving the illusion of phenomenality. We might call it pseudo-phenomenal consciousness. Access consciousness locks us directly onto the world, binding us to it in a complex web of sensitivities and reactions. Pseudo-phenomenal consciousness adds a new level of awareness, highlighting how we are locked onto the world and enabling us to respond to the interaction itself. It is a real process that does important work. We go wrong only when we take it too literally, mistaking figurative representations of complex physical properties for faithful representations of simple nonphysical ones. In *Consciousness Explained* (1991), Dennett draws a comparison with a computer's user interface, with its icons for files, folders, waste basket and so on. This is a fiction created for the benefit of the user (a 'user illusion'). By manipulating the icons, we can easily control the computer without knowing anything about its programming or hardware. Similarly, representations of phenomenal properties are simplified, schematic representations of the underlying reality, which we can use for the purposes of self-control. We should no more expect to find phenomenal properties in our brains than to find folders and waste baskets inside our laptops.

Why should you adopt the illusionist view? Well, to begin with, it offers a new approach to the problem of consciousness, which respects the intuitions of both sides. Illusionists agree with dualists that consciousness seems to have nonphysical features, and with physicalists that all the effects of consciousness can be explained in physical terms. By focusing on representations of phenomenal properties, it reconciles these claims. There are also more specific arguments for illusionism. I shall mention three.

The first concerns explanatory simplicity. If we observe something science can't explain, then the simplest hypothesis is that it's an illusion, especially if it can be observed only from one particular angle. This is exactly the case with phenomenal consciousness. Phenomenal properties cannot be explained in standard scientific ways and can be observed only from the first-person viewpoint (no one but me can experience my sensations). This does not show that they aren't real. It could be that we need to radically rethink our science but, as Dennett says

<https://www.ingentaconnect.com/contentone/imp/jcs/2016/00000023/f0020011/art00005> , the theory that they are illusory is the obvious default one.

A second argument concerns our awareness of phenomenal properties. We are aware of features of the natural world only if we have a sensory system that can detect them and generate representations of them for use by other mental systems. This applies equally to features of our own minds (which are parts of the natural world), and it would apply to phenomenal properties too, if they were real. We would need an introspective system that could detect them and produce representations of them. Without that, we would have no more awareness of our brains' phenomenal properties than we do of their magnetic properties. In short, if we were aware of phenomenal properties, it would be by virtue of having mental representations of them. But then it would make no difference whether these representations were accurate. Illusory representations would have the same effects as veridical ones. If introspection misrepresents us as having phenomenal properties then, subjectively, that's as good as actually having them. Since science indicates that our brains *don't* have phenomenal properties, the obvious inference is that our introspective representations of them are illusory.

There is also a specific argument

<https://www.ingentaconnect.com/content/imp/jcs/2018/00000025/f0020009/art00001> for preferring illusionism to property dualism. In general, if we can explain our beliefs about something without mentioning the thing itself, then we should discount the beliefs. For example, if we can explain people's beliefs about the paranormal in psychological terms, without mentioning actual paranormal events, then we've no reason to trust those beliefs. But if phenomenal consciousness is nonphysical, then the same goes for it. As I mentioned, there is good reason to think that all our mental processes, including the processes of belief formation, can be fully explained in physical terms. We would have exactly the same beliefs about phenomenal consciousness – that it is real, vivid and undeniable – even if we didn't really have it. So we shouldn't trust those beliefs.

By now, you are probably bursting with objections. I'll respond to some that I hear often. First, isn't consciousness the one thing that couldn't be an illusion? Isn't it the foundation of all our knowledge of the natural world? All we can be certain of in perception (the objection goes) is that we're having such-and-such experiences (say, of a reddish apple-shaped patch). We infer the existence of external objects from this.

This is a common view, but I think it gets things precisely the wrong way round. Think about it from an engineering perspective. If you were building an autonomous robot, you would begin by equipping it with sensors for significant features of the external

world and critical states of its own body, so that it could perform tasks, get the resources it needs, and protect itself from harm. Only later might you think of adding introspective systems so that it could monitor its own sensory processes and exercise sophisticated kinds of self-control.

It would be surprising if evolution hadn't taken the same path with us. There is, it is true, a sense in which pseudo-phenomenal consciousness is basic. We can detect that we are reacting as if there were an apple in front of us even when there is no apple, as when we hallucinate. But it doesn't follow that we infer the presence of real apples from our reactions. Because we introspectively monitor our perceptual processes, we tend to think that the monitoring is essential to perception, but it is doubtful that this is true. We often perform complex tasks, such as driving, without attending to what we are doing. It is also true that we find it hard to doubt what introspection tells us. But this might be because there is no easy way of checking its accuracy. It doesn't follow that it is infallible. (We have the final word about our conscious experiences, but it might not be the right word.) In fact, there is reason to think that introspection is unreliable. For example, it can be shown experimentally that we have far less awareness of objects in the periphery of our vision than we think we do. And, of course, if we could observe our own brain processes externally, as a neuroscientist does, we wouldn't observe any phenomenal properties.

It is tempting to suppose that there is a boss system, a self, to which all other mental systems report

Second, isn't the very idea

<<https://www.ingentaconnect.com/contentone/imp/jcs/2016/00000023/f0020011/art00010>> of illusionism confused? To be under the illusion of seeing an apple is to have an experience exactly like that of seeing an apple, even though there's no apple present. How then could we be under the illusion of having an experience? If you are having an experience exactly like a pain experience, then you are having a pain experience. As the philosopher John Searle puts it in *The Mystery of Consciousness* (1997), when it comes to consciousness, the appearance is the reality. This looks like a serious objection, but in fact it is easily dealt with. Properties of experiences themselves cannot be illusory in the sense described, but they can be illusory in a very similar one. When illusionists say that phenomenal properties are illusory, they mean that we have introspective representations like those that we would have if our experiences had phenomenal properties. And we can have such representations even if our experiences don't have phenomenal properties. Of course, this assumes that the representations themselves don't have phenomenal properties. But, as I noted,

representations needn't possess the properties they represent. Representations of redness needn't be red, and representations of phenomenal properties needn't be phenomenal.

But how does a brain state represent a phenomenal property? This is a tough question. I think the answer should focus on the state's effects. A brain state represents a certain property if it causes thoughts and reactions that would be appropriate if the property were present. I won't try to develop this answer here. For it is not only illusionists who must address this problem. The notion of mental representation is a central one in modern cognitive science, and explaining how the brain represents things is a task on which all sides are engaged. Indeed, even realists about phenomenal consciousness must explain how we mentally represent phenomenal properties, if they are to account for the fact that we think and talk about them. There is a challenge here for illusionism but not an objection.

Finally, who is the subject of this illusion? Doesn't illusionism presuppose a conscious subject who experiences the illusion? My answer is that the subject is the person as a whole, the autonomous evolved organism composed of interacting biological subsystems. 'We' are aware of something if information about it reaches enough of our neural subsystems for us to be able to think and act flexibly with respect to it – to use it, remember it, tell others about it and so on. Think of a large organisation composed of many departments, each responsible for one function but sharing information with each other. If enough departments possess and use a certain piece of information, then the organisation as a whole can be said to be aware of it. The same goes for biological organisms such as us. If enough mental systems receive and use representations of a certain property, then the organism itself can be said to be aware of the property. And if the representations are illusory, then the organism is under an illusion. It is tempting to suppose that there is a boss system, a self, to which all other mental systems report, and that we are aware of something only if the boss system gets to know about it. But I would argue that this boss system is itself an illusion. That is another story, however.

The subjective world of phenomenal consciousness is a fiction written by our brains in order to help us track the impact that the world makes on us. To call it a fiction is not to disparage it. Fictions can be wonderful, life-enhancing things that reveal deep truths about the world and can be more compelling than reality. Unlike Neo in *The Matrix*, you shouldn't want to escape this fictional world; it's a benign one, designed by evolutionary processes to help you thrive. But you shouldn't mistake it for reality either.

Acknowledgement: the author thanks Daniel Dennett for his comments on an earlier draft of this article.

Keith Frankish is a philosopher and writer. He is an honorary reader in philosophy at the University of Sheffield, a visiting research fellow with the Open University, and an adjunct professor with the Brain and Mind programme at the University of Crete. He lives in Greece.

aeon.co26 September, 2019