



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/212812/>

Version: Published Version

Article:

Frankish, K. (2024) The ethical implications of illusionism. *Neuroethics*, 17. 28. ISSN: 1874-5490

<https://doi.org/10.1007/s12152-024-09562-5>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



The Ethical Implications of Illusionism

Keith Frankish 

Received: 15 February 2024 / Accepted: 2 May 2024
© The Author(s) 2024

Abstract Illusionism is a revisionary view of consciousness, which denies the existence of the phenomenal properties traditionally thought to render experience conscious. The view has theoretical attractions, but some think it also has objectionable ethical implications. They take illusionists to be denying the existence of consciousness itself, or at least of the thing that gives consciousness its ethical value, and thus as undermining our established ethical attitudes. This article responds to this objection. I argue that, properly understood, illusionism neither denies the existence of consciousness nor entails that consciousness does not ground ethical value. It merely offers a different account of *what* consciousness is and *why* it grounds ethical value. The article goes on to argue that the theoretical revision proposed by illusionists does have some indirect implications for our ethical attitudes but that these are wholly attractive and progressive ones. The illusionist perspective on consciousness promises to make ethical decision making easier and to extend the scope of our ethical concern. Illusionism is good news.

Keywords Consciousness · Illusionism · Ethics · Value · Animal consciousness

The illusion is irresistible. behind every face there is a self. We see the signal of consciousness in a gleaming eye and imagine some ethereal space beneath the vault of the skull lit by shifting patterns of feeling and thought, charged with intention. an *essence*. but what do we find in that space behind the face, when we look?. The brute fact is there is nothing but material substance: flesh and blood and bone and brain. I know, I've seen. You look down into an open head, watching the brain pulsate, watching the surgeon tug and probe, and you understand with absolute conviction that there is nothing more to it. There's no one there. It's a kind of liberation.— Paul Broks ([1], p. 17)

Introduction

Illusionism is a revisionary view of consciousness, which denies the existence of the “phenomenal” properties traditionally thought to render experience conscious. The view has theoretical attractions, but some think that it also has objectionable ethical implications. They take illusionists to be denying the existence of consciousness itself, or at least of the thing that gives consciousness its ethical value, and thus as undermining our established ethical attitudes. This should give illusionists pause. Though ethically troubling views might be true, we should be cautious

K. Frankish (✉)
Department of Philosophy, The University of Sheffield,
Sheffield, UK
e-mail: k.frankish@gmail.com

in advocating them, especially as we might be wrong about their truth.

This article responds to this concern about illusionism, arguing that the ethical objections beg the question against the view. Properly understood, illusionism neither denies the existence of consciousness nor entails that consciousness does not ground ethical value. It merely offers a different account of *what* consciousness is and *why* it grounds ethical value, and it no more undermines our ethical attitudes than other revisions to our view of human nature, such as the denial of a vital force. It is true that the theoretical revision proposed by illusionists does have indirect implications for our ethical practice, but these are attractive and progressive ones. The illusionist perspective on consciousness promises to make ethical decision making easier and to extend the scope of our ethical concern, and, far from being incompatible with a humanistic conception of ourselves, it can in fact enrich and extend it. Illusionism is good news.¹

The article is structured as follows. The section "Illusionism" briefly introduces illusionism and the reasons for endorsing it. The section "Does Illusionism have Negative Ethical Implications?" outlines and responds to various reasons for thinking that illusionism has negative ethical implications. The section "Positive Ethical Implications of Illusionism" then sketches some positive ethical implications of illusionism, arguing that the view enables a more progressive, less anthropocentric ethics of consciousness. A brief conclusion rounds things off.

For convenience, I shall often speak about what *illusionists* claim and how they respond to objections. In doing this, I do not mean to stipulate what illusionism is or to speak for all illusionist-aligned theorists. Illusionism is a broad church, and the basic position can be developed and defended in many ways. For most of the time, however, I shall be relying on general claims that are central to the illusionist approach, and all the points I make should be at least *options* for illusionists. When I invoke my own more specific views, I shall make this clear.

¹ Since illusionism is in part a response to progress in neuroscience, this article aligns with the optimistic, anti-essentialist materialism advocated by Mette Høeg in her paper in this issue.

Illusionism

Phenomenal Realism

Illusionism is not so much a theory as a theoretical approach — a way of conceptualizing consciousness and an agenda for theorizing about it. It is not a new approach although the name is relatively new.²

The easiest way to introduce illusionism is by contrast with a more traditional approach, which I shall call *phenomenal realism*. Presentations of phenomenal realism often take their start from Thomas Nagel's famous remark that an organism has conscious mental states, "if there is something that it is like to *be* that organism — something it is like *for* the organism" ([14], p. 436). Nagel calls this aspect the *subjective character* of experience, and he claims that it cannot be analysed in functional or intentional terms or inferred from the physical features of the experiencing organism:

if the facts of experience — facts about what it is like *for* the experiencing organism — are accessible only from one point of view, then it is a mystery how the true character of experiences could be revealed in the physical operation of that organism. ([14], p. 442)

Nagel does not deny that conscious mental states have a basis in brain states that play distinctive functional roles. His claim is that they also have a subjective aspect that cannot be analysed in functional terms. The subjective character of experience exists only *for* the subject (Nagel borrows Sartre's notion of the *pour soi*), and it constitutes an "internal world" ([14], p. 445) inaccessible to public investigation. I shall express this by saying that this aspect of experience is *irreducibly subjective*. Nagel argues that our only way of understanding what others' experiences are like is by imaginative extrapolation from our own and that the internal worlds of creatures markedly different from ourselves, such as bats, are forever closed to us.

Nagel's way of defining consciousness has been hugely influential. Although hard to make precise,

² For defences of broadly illusionist positions, see, e.g., [2–12]. For a more detailed introduction to illusionism and further references, see [13].

the “what-it-is-like” formulation is widely felt to be compelling, and it is the starting point for much theorizing about consciousness, including for what some regard as one of the leading scientific theories of consciousness, Integrated Information Theory [15]. The irreducibly subjective what-it-is-like properties of a conscious mental state are often referred to as its *phenomenal properties*, or *qualia*, and by *phenomenal realism* I mean the view that conscious experiences have phenomenal properties.³

Phenomenal realism opens a rift in our conception of the mental — a rift between the objective side of the mind, which can be studied by the cognitive sciences, and an essentially subjective side, which is known only to the subject whose mind it is. The objective side is often conceptualized in functional terms, with mental states defined by the role they play in the processes of bodily and behavioural regulation, and I shall adopt that approach here.⁴ I shall also assume that in so far as we are aware of such states in ourselves, it is by the mediation of functional representational processes. The subjective side, by contrast, is constituted by irreducibly subjective phenomenal properties, presented to the self like a show in a mental theatre (Daniel Dennett calls it the *Cartesian Theatre* [3]). Phenomenal realists typically hold that these properties are known to their possessor by *direct acquaintance* — an unmediated form of awareness, which reveals their essence (e.g., [16]).

Take pain, for example. In the functional sense, pain is a hugely complex reaction to harmful stimuli, which involves a host of physiological, psychological,

and behavioural changes. In the phenomenal sense, pain is simply a subjective awfulness, an immediately known mental quality, distinct from all reactions. A similar distinction can be drawn with respect to consciousness itself. In the functional sense, a mental state is conscious if it has a certain functional role — say, if it has appropriate effects on a wide range of cognitive processes. It is phenomenally conscious if it possesses a what-it-is-likeness. Chalmers refers to these two concepts of the mental as *psychological* and *phenomenal* respectively ([17], ch.1), and when I speak of *psychological states*, I shall mean mental states in the functional sense.⁵

Phenomenal realism has many problems. First, there is the notorious “hard problem” of explaining how brain states come to have phenomenal properties [18]. There seems no chance of reductively explaining the phenomenal in terms of the physical. No account of objective processes would entail the existence of an irreducibly subjective world associated with them. The best the phenomenal realist can do is to propose metaphysical theories of how physical states and phenomenal states are *correlated*. A wide range of theories have been proposed, with wildly different patterns of correlation. Some theories restrict phenomenal consciousness to humans [19], or to biological organisms [20, 21], or to systems with a certain informational structure [22]; but others ascribe it much more liberally — to fundamental particles [16, 23], to every arrangement of fundamental particles [24, 25], and to the universe as a whole [16, 26]. Such theories can be assessed for various theoretical virtues, but there is no way of empirically confirming that the posited correlations hold. The best we can do is to establish correlations between physical states and psychological or behavioural responses we take to be *symptomatic* of consciousness. Moreover, as Dennett has shown, even first-person knowledge of phenomenal properties is riddled with problems, since memory plays an ineliminable role in all our judgements about them [2].

These problems for phenomenal realism may not be conclusive, but they are sufficient to motivate

³ In other presentations of illusionism, I have focused on the claim that phenomenal properties are *anomalous* [7] or *not functionally characterizable* [13], but I have come to think that irreducible subjectivity is the heart of the matter, and for present purposes I shall treat a commitment to irreducibly subjective mental properties as the core of phenomenal realism, and its denial as the core of illusionism.

⁴ Note that I use “functional” in a broad and basic sense; a functional state is one defined by what it *does* — the role it plays. There is no assumption that functional states must be characterized at a large scale, abstracting away from the micro-level details of neuroanatomy and neurophysiology. Thus, a functionalist approach does not imply that micro-level features do not matter to the explanation of mental phenomena (on the contrary, we know they matter *a lot*). It claims merely that they matter because of what they *do* — the functions they perform — not because of intrinsic features that do not make a causal difference. For discussion, see [4], pp. 17–22.

⁵ The functional/psychological form of consciousness is often referred to as “access consciousness”, and the present paper can be seen as exploring the implications of denying that there is a phenomenal form of consciousness in addition to the access form.

exploration of an alternative. The standard physicalist approach is to argue that phenomenal properties are identical with physical ones and that our intuitions to the contrary can be explained by special features of the way we conceptualize phenomenal properties (the “phenomenal concept strategy”) [27–29]. A huge amount of work has been done on this approach, which I shall not attempt to critique here. I shall merely note that if phenomenal properties are irreducibly subjective, then it is hard to see how the strategy can work. Its defenders must either explain how physical properties can be accessible only from one point of view or deny that phenomenal properties are irreducibly subjective. And if they take the latter option, then they must explain how their form of phenomenal realism differs from the illusionist approach described next.

The Illusionist Alternative

Illusionism involves the rejection of phenomenal realism. Illusionists deny that consciousness involves awareness of irreducibly subjective properties, revealed directly to the self. Different illusionist theorists propose different positive accounts of consciousness, but they all agree that conscious experiences can be fully described in broadly functional, third-person terms, that they can be reductively explained in terms of embodied, environmentally embedded neural processes, and that our knowledge of them is mediated by representational mechanisms.

Illusionists do not deny that we have a strong conviction that our experiences have an irreducibly subjective aspect. When we attend to our own experiences (when we *introspect*), it can seem natural to think of them as having an intrinsic what-it-is-likeness distinct from all the effects they have. However, illusionists say that this is a sort of illusion created by the brain’s processes of self-monitoring and self-modelling. These processes track functional aspects of experience but misrepresent them in a way that induces us to believe that we are directly acquainted with irreducibly subjective properties. In fact, such properties exist only as intentional objects of our introspective attitudes. If introspection is an evolved

representational process, then the possibility of such systematic misrepresentation must be open.⁶

Illusionists differ in their accounts of the nature of the introspective monitoring involved and how it generates belief in phenomenal properties, but they typically hold that some aspect of the process is cognitively impenetrable (or at least strongly resistant to cognitive penetration), so that a disposition to believe in phenomenal realism persists even when the view itself is explicitly rejected. It is this resistance to correction that makes it appropriate to talk of illusion here.⁷

The illusionist programme thus has two aims: to provide a positive account of conscious experience in broadly functional terms and to explain why we are inclined to judge that conscious experiences have an irreducibly subjective nonfunctional aspect. I have called the second question the *illusion problem* [7].

It is crucial to stress that in denying that conscious experiences are *irreducibly* subjective, illusionists are not denying that they are subjective in *any* sense. Far from it. Each organism has a subjective *perspective* on the world, thanks to a set of perceptual capacities crafted by evolution and tuned by experience to track features important to its well-being. And each organism exhibits a multifaceted subjective *response* to the features it tracks — a huge cluster of neurally modulated physiological, psychological, and behavioural reactions, shaped again by evolution and personal experience. Different organisms are sensitive to different features and have different responses to the same features, giving each its own subjective *take* on

⁶ Because of its appeal to introspective monitoring, illusionism might be thought to be a variant of higher-order representation (HOR) theories of consciousness [30]. This would be a misunderstanding, however. HOR theorists are typically phenomenal realists, who believe that higher-order representations are necessary for phenomenal consciousness. Illusionists, by contrast, hold that consciousness consists in the first-order functional processes being monitored. The monitoring does not confer a further phenomenal aspect upon these processes; it merely disposes us to believe that they have such a further aspect.

⁷ Of course, the term “illusion” is being used here in a functional sense, not a phenomenal one. To be under the illusion of perceiving or introspecting something is to undergo the psychological reactions one would undergo if one really were perceiving or introspecting it, including being disposed to believe that it is real.

the world. This subjective take is perfectly real; it is just not *irreducibly* subjective.

The version of illusionism I prefer treats the response patterns associated with perceived features as playing a central role in creating the sense that we have an irreducibly subjective inner life [31].⁸ Each stimulus produces a huge wave of neural adjustments to systems controlling attention, expectation, motivation, emotion, and many other functions, adjusting the organism's global state and orientation to the world. This pattern of responses embodies the *significance* the perceived feature has for the organism, and information about it will be hugely useful to reflective, social creatures like us, allowing them to adopt deliberate strategies for cultivating or avoiding experiences according to their significance. I suspect that this led to the evolution of brain systems dedicated to monitoring response patterns and producing simplified, schematic models of them, which feed into higher-level cognitive processes, giving us the sense that stimuli produce simple, intrinsically meaningful mental qualities in us. Dennett speaks of this as a *user illusion*, likening it to the graphical interface on a personal computer, which allows its user to manipulate complex data in a simple and intuitive way [3, 34]. In a similar way, the brain's user illusion allows us to detect, control, and communicate the immensely complex psychological responses that constitute our subjective take on the world. It is highly useful — provided we don't mistake it for an insight into fundamental reality.

Does Illusionism have Negative Ethical Implications?

Does Illusionism Deny Experience?

The basic ethical objection to illusionism is that it denies the existence of something that grounds ethical value, perhaps the only thing that grounds it — conscious experience. Galen Strawson gives forthright expression to the worry. He labels denial of

phenomenal properties “the Denial” and has this to say about it:

It's important to be clear that there is no suffering if the Denial is true. [. . .] There's no joy either, no feeling at all. But what can sometimes seem most important is that there is no suffering — in spite of clinical depression and thousands of other extraordinarily painful diseases, murder, rape, famine, slavery, bereavement, torture, genocide. What is more, no one has ever really caused anyone any pain. ([35], p. 36)

But this is a highly uncharitable reading of what illusionists say. Illusionists do not claim that *conscious experiences* are illusory — that we cannot really see, hear, feel pain, and so on. What is illusory, according to the illusionist, are the *phenomenal properties* of conscious experience — the irreducibly subjective aspect, the private show in the Cartesian Theatre. They are not denying conscious experience but claiming that we have a mistaken conception of what conscious experience *is* — a conception we find compelling because of the way our introspective systems work. As Dennett says:

I don't deny the existence of consciousness; *of course*, consciousness exists; it just isn't what most people think it is, as I have said many times. ([36])

Such reconceptualizations are common in the history of science. When Hippocrates proposed that the “sacred disease” was not in fact a divine affliction, he was not callously denying that people ever had epileptic seizures, and when twentieth-century biologists denied that life was a vital force they were not claiming that everyone was dead. Similarly, when illusionists claim that pain is not an irreducibly subjective state, they are not denying the reality of suffering. Indeed, in offering an account of what pain *really is*, they are assuming that pain *does* exist; if it didn't, illusionism would be *false*.

Now, phenomenal realists can and do respond that this is just playing with words — attaching the name “consciousness” to something else. Here is Strawson again:

when Dennett says that consciousness exists he *reversifies* or *looking-glasses* the ordinary meaning of the word “consciousness”. That is,

⁸ The role of reactions in generating our sense of phenomenality is a central theme in work on consciousness by Daniel Dennett [2–4] and Nicholas Humphrey [9, 32, 33].

he uses the word in such a way that what he means by it *excludes what the word actually means*. More moderately: he uses the word in such a way that what he means by it excludes what it is standardly (almost universally) used to mean — especially in discussions of this sort. ([35], p. 33)

What people standardly mean by “consciousness” is a matter for empirical investigation. Maybe many do think of consciousness as an irreducibly subjective state. Illusionists don’t deny that. (Indeed, they claim that we are psychologically disposed to think that way; that’s why they talk of illusion.) But it doesn’t follow that people are *right* to conceptualize consciousness as irreducibly subjective, any more than their ancestors were right to conceptualize epilepsy as a divine affliction.

Of course, this response works only if both parties are *referring* to the same thing when they talk of a conscious state — picking out the same worldly phenomenon despite their different conceptions of its nature. People were able to disagree about whether epilepsy was a sacred disease or a form of brain dysfunction because they could agree in identifying paradigm cases of the condition. And phenomenal realists will say that illusionists are *not* referring to the same thing they are.

Now, if phenomenal realists stipulate that “conscious state” means an irreducibly subjective state, then it is true that illusionists aren’t talking about that. But this is no more persuasive than stipulating that “epilepsy” means a disease with a divine origin. We can’t resolve disputes about the nature of things by stipulative definition. A better way to proceed is by pointing to examples. For example, suppose that someone punches me on the nose — hoping, perhaps, that this will convince me of the truth of phenomenal realism. I enter a distinctive state, which I dislike very much. This state is one I have learned to call “pain”, and it is an example of the kind of state we’re talking about. We can tell when we are undergoing such states, and we can often tell to a moral certainty when other creatures are undergoing them. (No one but a philosopher could seriously doubt that an injured, whimpering dog is in pain.) Being able to detect these states is a skill, which does not require explicit theorising about the nature of the states or the means by which we detect them. We can point to endless

further examples: the states we enter when looking at a ripe banana, smelling gorgonzola cheese, tasting Marmite, and so on. These are the states we are talking about. I shall refer to them as *paradigm conscious states* (*paradigm experiences*, *paradigm pains*, etc.). The phenomenal realist says that paradigm conscious states have an irreducibly subjective aspect; the illusionist denies it.⁹

This is the most natural way of picking out paradigm experiences, but an alternative method, which should come to the same thing, is to piggyback on the phenomenal realist’s own definition. Illusionists can say that paradigm conscious states are the ones we *describe* as being “like something” to undergo. The illusionist doesn’t deny that there are such states, and they are certainly the ones the realist is talking about. Again, the question is whether these states have an irreducibly subjective aspect.

Of course, the illusionist answer to that question may be wrong. Illusionists may have a deeply flawed view of what paradigm pain is; but they do not deny its reality, and their position does not have the dire implications that such a denial would have.

Does Illusionism Deny the Value of Conscious Experience?

Phenomenal realists will doubtless reply that even if illusionism does not deny consciousness outright, it denies the *aspect* of consciousness that makes it ethically valuable. They will say that phenomenal properties, especially valenced ones, such as pain and pleasure, ground an intrinsic, noninstrumental form of ethical value, and that to conceptualise consciousness in functional terms is to deny this source of value. Functionally conceived (they will say), pain is just a causal state that has either no intrinsic value, or, at any rate, considerably less than the phenomenal form.

François Kammerer — himself an illusionist — claims that this presents a *normative challenge* for illusionists [37]. He argues that there is a strong case for thinking that phenomenal properties *would* ground intrinsic value if they existed. He calls this

⁹ The dialectic here is that Strawson accepts that illusionism would entail scepticism about consciousness, and I am running a paradigm case argument against him.

thesis *Phenomenal Value*, and in support of it he asks us to imagine a case of what he calls *Pure Suffering*:

Piotr, a genuinely conscious being endowed with phenomenal states, suffers from a rare disease. Every night, during sleep, he wakes up and, for one hour, feels the most excruciating pain one can imagine. The pain is so intense that it leaves him unable to move or to talk (or to have any observable behavior). For one hour, his life is entirely filled by nothing but the most awful pain ever felt [. . .] He then goes back to sleep. In the morning, he systematically forgets about everything that has happened to him during the night: these nocturnal episodes of pain do not leave any memory, conscious or unconscious, and do not change his psychological dispositions. ([37], p. 899)

Kammerer takes it to be intuitive that such *Pure Suffering* has negative value and that it has it in virtue of its phenomenal aspects. (The example is designed to reduce the nonphenomenal candidates for grounding negative value, such as traumatic memories.) Even illusionists can accept this, Kammerer argues. For though they deny that *Pure Suffering* occurs, they can agree that it *would have* negative value if it did. Kammerer further suggests that this commitment to *Phenomenal Value* underwrites some of our important ethical attitudes, including the belief that conscious mental states matter more than nonconscious ones and the lives of sentient beings more than those of nonsentient ones. The normative challenge for illusionists is to say what normative revisions are required by the denial that this putative source of value exists.¹⁰

Kammerer analyses the options available to the illusionist and argues that illusionists will probably have to accept some revisionary ethical consequences, at least if they endorse *Phenomenal Value*. He asks us to imagine Piotr's zombie twin, who every night undergoes *Pure Zombie Suffering*, defined as.

a state that is physically and functionally indistinguishable from the state of Piotr when Piotr feels the most excruciating pain one can imag-

ine. For this one hour, Zombie Piotr is unable to move or to talk (or to have any observable behavior), in virtue of the functional properties of the state he is in. Even though he does not have any phenomenal experience, all of his functionally conceived mental states — desires, judgments, etc. — are similar to the ones had by Piotr when Piotr suffers ([37], p. 905)

Kammerer accepts that Zombie Piotr's form of suffering has negative value but argues that, if *Phenomenal Value* is true, then it is very unlikely to have *exactly the same* negative value as Piotr's. For that would mean that the negative value of Piotr's suffering would have two distinct but equivalent grounds: the phenomenal properties instantiated only in Piotr's case and the nonphenomenal ones instantiated in both cases. And that, Kammerer claims, would be an "amazing coincidence" ([37], p. 906).

Kammerer's articulation of the normative challenge is useful, but I think it gives too much weight to phenomenal realist intuitions. After all, if illusionism is true, then no one has ever undergone genuinely phenomenal pain. So what exactly are we imagining when we imagine Piotr's *Pure Suffering*? Speaking for myself, I try to recall the most awful pain I have ever experienced and then imagine Piotr being in a state like *that* or even worse. Then I add in thought that the state I'm imagining is an irreducibly subjective one that can't be characterized in functional terms. That's the best I can do. It's like being asked to imagine a creature that is animated by a vital force. I just imagine some living thing, while telling myself that its vital processes depend on a mysterious force. In both cases, I imagine a familiar thing with a certain theoretical gloss. And of course the result is that I judge Piotr's suffering to have negative value, since the state I recall was an extremely distressing one.

But if illusionism is true, then the state I am recalling was in fact a complex functional one, and when I imagine Piotr being in *that* state, I am imagining him being in a state that is in fact a complex functional state, even though I conceptualize it otherwise. And when I imagine Zombie Piotr's state, I should be imagining *exactly the same state*, though without conceptualizing it as something it is not.

So the only difference between the scenarios is in the theoretical gloss I add, and this should make little difference to their relative ethical value. I may regard

¹⁰ As Kammerer presents it, the challenge assumes that normative judgments have truth values, and I shall accept this assumption for present purposes.

Piotr's state as harder to explain than that of his zombie twin, but this won't give it greater negative value. (If anything, it might give it a more positive aspect, in virtue of its metaphysical specialness.) What I find distressing about pain isn't that it poses a hard problem, and, sadly, belief in illusionism doesn't make pain any less distressing. Besides, we can assume that Piotr and his zombie twin both conceptualize their state in the same way, as genuinely phenomenal (conceptualizing being understood as a functional process).

I grant that we have an intuition that Zombie Piotr's state isn't as bad as Piotr's. But that is because we are invited to imagine it in a very different way — not by recalling an episode of paradigm pain but by imagining the functional structure of paradigm pain. However, this structure will be hugely complex, and we have only the sketchiest idea of what it is. So, what we imagine is a handful of large-scale functional states (“desires, judgements, etc.”), and we conclude, rightly, that what we are imagining has less negative value than paradigm pain. However, if we really could imagine Zombie Piotr's functional state in all its detail and complexity, then — if illusionism is true — we would be imagining a state identical to the one we imagined Piotr being in, minus the theoretical gloss. Thus, far from being an amazing coincidence, it will be a triviality that they have exactly the same value, and no revisionary consequences will be dictated.

Kammerer might object that I am not imagining Piotr's state correctly. I should be imagining a case of distinctively *phenomenal* pain — pain that is the way it introspectively seems to be, not pain as it actually is. To imagine *Pure Suffering*, we have to imagine pain that really does have irreducibly subjective phenomenal properties. If illusionism is true, this will be different from imagining paradigm pain, since paradigm pain lacks such properties.

This is true, but it is doubtful that the difference would be detectable *by the subject*. For if illusionism is true, paradigm pain is introspectively *represented as* phenomenal, and both Piotr and Zombie Piotr will have exactly the same set of introspective representations. And though one set will be veridical and the other not, this difference won't be introspectively detectable, and the two sets will have the same psychological effects. Piotr and his zombie twin will also possess the same exteroceptive and interoceptive

representations, so if Piotr detects any other effects of his phenomenal pain, Zombie Piotr will represent himself as detecting them too.

Again, Kammerer may object that I am not imagining Piotr's state correctly. For I am supposing that awareness of phenomenal pain would be constituted by representational processes of a functional kind, which would be shared with his zombie twin. But — the objection goes — if phenomenal pain were real, our awareness of it would not be constituted in this way. We would have direct *acquaintance* with it, independently of all functional introspective processes. But again, how are we to imagine this? If illusionism is true, we have never been directly acquainted with phenomenal pain; we have only ever had distorting, functionally mediated awareness of psychological pain, which we falsely believe to be direct acquaintance with phenomenal pain. Would genuinely direct acquaintance with phenomenal pain be worse than this? I have no idea.

There is a further problem with *Phenomenal Value*. The cases of Piotr and Zombie Piotr pump the intuition that pure phenomenal pain would be worse than pure functional pain. But if phenomenal properties are irreducibly subjective, then it is arguable that the opposite is true.¹¹ For paradigm pain *involves* functional processes. Only an idealist would deny this. And since these functional processes are *not* irreducibly subjective, they won't be part of phenomenal pain. Thus, to assess the value of *pure* phenomenal pain, we need to *subtract* all the functional aspects of paradigm pain that might contribute to its negative value. Kammerer gestures at this by supposing that Piotr is immobilized by the pain and that he forgets the episode as soon as it ends. But that removes only overt pain behaviour and the lasting effects of pain. To properly imagine pure phenomenal pain, we have to imagine that during his nighttime episodes Piotr undergoes none of the *psychological* reactions typically involved in paradigm pain. We have to imagine that he is not uncomfortable, stressed, or tensed up, isn't aware of a racing heart or an urge to vomit, doesn't have his attention riveted on the condition of some part of his body, doesn't find it difficult to think

¹¹ I should note that Kammerer does not specify that phenomenal pain is to be understood as irreducibly subjective, so he may understand the Piotr scenario somewhat differently.

calmly or to entertain agreeable thoughts, doesn't desire relief, isn't anxious or fearful, and does not even believe that he is in pain (all these states being conceptualized functionally). In short, we have to imagine that he is in a condition that is psychologically identical to a state of complete painlessness *but which is still excruciatingly painful*. I don't think I can do this, and I certainly don't have the intuition that the condition would be worse than the inverse one of undergoing all the psychological aspects of paradigm pain without any of the phenomenal ones. Just the opposite.¹²

In short, our intuition that phenomenal pain would be worse than functional pain exists only because illusionism isn't taken seriously as an account of paradigm pain. Once the illusionist alternative is on the table, there is no reason to trust the intuition or to think that illusionism dictates a revision to our view of how bad pain is.¹³

Can Illusionism Explain why Pain is Bad?

Kammerer argues that illusionists face a further challenge. They must explain *what it is* that makes pain bad and *justify* our belief that it has this bad-making property ([37], p. 907). Phenomenal realists will say that pain is bad because it possesses an intrinsic phenomenal nastiness and that we are justified in believing that it has this property because introspection fully reveals its nature to us. What can illusionists say?

This is an interesting challenge, but I don't think our everyday ethical practices presume a specific answer to it. Just as we can all agree on identifying paradigm cases of pain, so we can all agree that pains are paradigm examples of intrinsically bad states. It is this agreement that guides our ethical practice, and it is compatible with a range of theoretical views about what grounds the badness of pain and what justifies our judgements about it. (By an *intrinsically*

bad state, I mean a state that is bad in itself, rather than because of some other state that it promotes. An intrinsically bad state in this sense could be a functional state of the thing that possesses it.)

This isn't to say that illusionists needn't address the challenge (Kammerer outlines some options they might take). A proper treatment is beyond the scope of this paper, but I shall sketch the approach I favour. The short answer is that pain states are bad because of how they *affect* us. When in severe pain, we have a powerful sense that some region of our body is being harmed. This sense obtrudes itself upon us, seizing our attention, disturbing our patterns of thought and activity, and creating an overwhelming desire for relief. In pain, we are passive, helpless, restless, anxious, and fearful. It is a highly aversive state, which we fear, strive to avoid, and are concerned to witness in others.¹⁴ Pain's badness lies precisely in such *reactions*. This of course reflects the evolutionary function of pain, which is to enhance our fitness by teaching us to avoid or escape situations that are harmful to our well-being. In creatures like us, who are capable of self-reflection, such aversive states reliably produce the belief that we are in a bad state, and such beliefs are justified by their reliable connection with the bad-making features.

Of course, people in pain don't articulate this line of thought, analysing the attentional, cognitive, conative, and affective components of the state that make it so aversive. They simply say that it *feels bad*. But this reflects a lack of knowledge of what pain really is, rather than a deep insight into its nature. The processes that generate our introspective judgements are unconscious, and saying that we know that pain is bad because we are directly acquainted with its badness amounts to little more than saying that we can't give a substantive answer.¹⁵

In short, conceptualizing pain as a complex set of psychological and physiological reactions, rather than

¹² The position sketched here is close to the one Kammerer calls "Scepticism about our value intuitions given our realist intuitions", which involves denying *Phenomenal Value* ([37], p. 917).

¹³ Leonard Dung develops a closely related line of reply to Kammerer in the course of defending illusionism against the charge that it entails scepticism about animal consciousness [38].

¹⁴ All the mental states mentioned here are, of course, to be understood as functional ones, whose putative phenomenal aspects yield to illusionist treatment.

¹⁵ Daniel Shabasson has argued that we unconsciously — and erroneously — infer the presence of phenomenal properties in order to provide a justificatory basis for our introspective judgements [39]. If he is right, then phenomenal properties are a sort of fiction we unconsciously invent in order to provide us with answers to questions such as "How do you know the banana seems yellow?" and "How do you know pain is bad?"

a phenomenal essence, does not make pain less bad. If anything, it makes it more real and more potent. The same, of course, goes for positively valenced states, such as joy.

Does Illusionism Eliminate Phenomenal Concepts?

Another worry about illusionism, outlined by Katalin Balog, is that it undermines our conception of ourselves and our worth [40]. Balog takes illusionists to be proposing that we eliminate phenomenal concepts (concepts of what experience is like) and reconceptualize ourselves in third-person terms, as “unconscious, but very intelligent robots: unfeeling, incapable of an inner life, not fully rising to personhood” ([40], p. 17) — a change which, she rightly notes, would undermine belief in human moral worth.

Balog is a physicalist, who employs the phenomenal concept strategy to defend a form of phenomenal realism.¹⁶ I shall not address her defence of that view here (for some relevant remarks, see my [41]), but merely comment on the illusionist attitude to phenomenal concepts.

It should be clear from what I have already said that illusionists do not make the radical recommendations Balog describes. They do not suggest that we are unconscious or that we stop talking about what our experiences are like. Far from it. In claiming that talk about what experience is like does not track irreducibly subjective properties, they are not claiming that it tracks *nothing at all*. (Think again of Hippocrates or the anti-vitalists.) As I explained earlier, I believe that such talk tracks something of huge importance: the subjective significance stimuli have for us, as embodied in our psychological reactions to them. The illusion is that it tracks simple, irreducibly subjective mental qualities, rather than these complex patterns of effects. Calling this an illusion isn't to denigrate it. The user illusion created by introspection is immensely useful, enabling a new level of self-description and self-control, and it is central to our conception of ourselves. Its value (like that of all designed illusions) lies precisely in its effects — in how it moves and motivates us. Illusionists do

not propose that we replace it with a neuroscientific image of ourselves, any more than they propose that we replace the graphical interfaces on our laptops with command-line ones. They merely urge us to recognize that it *is* a kind of illusion and not a window onto a metaphysically private world that defies scientific explanation. Once this point is made, I think Balog's fears should fall away.

Positive Ethical Implications of Illusionism

We have seen that illusionism should not make us regard paradigm experiences as less ethically significant than we ordinarily take them to be, and to that extent does not dictate any revisions to our ethical attitudes. However, the view may still have indirect ethical implications. The illusionist reconceptualization of consciousness has consequences for how we study consciousness and for our judgements about which beings possess it. (Compare Hippocrates' reconceptualization of epilepsy. Thinking of epilepsy as a brain disfunction rather than a divine affliction didn't make epilepsy any less distressing, but it did affect how physicians treated it and which creatures they thought could suffer from it.) And this in turn may have implications for our ethical practices. I believe that illusionism does have such implications, and that they are positive ones. I shall discuss them in this section.

Knowledge of Consciousness

The first implication concerns our knowledge of consciousness in others, particularly in nonhumans. If phenomenal realism were true, then we would have no direct access to facts about what other creatures' experiences are like. Phenomenal properties (in the Nagelian sense that is our focus here) are irreducibly subjective, and all we could do is look for objective features we take to be correlated with them, such as reports, reactions, informational structures, or patterns of neural activity. But, as noted earlier, there is deep disagreement about *which* features are correlated with phenomenal consciousness, with the options ranging from subatomic particles to the entire universe. And there is no experimental way of testing the different accounts. The best we can do is to compare them for theoretical virtues, such as elegance

¹⁶ I am not sure whether Balog would endorse the form of phenomenal realism discussed here, which is committed to the existence of irreducibly subjective mental properties.

and economy. And even here, intuitions vary widely. Is it more economical to suppose that phenomenal consciousness is restricted to animals like ourselves, who possess paradigm experiences, or to suppose that consciousness is a fundamental property of all matter? Intuitions vary!¹⁷

It might be suggested that we could look for causal effects that indicate the presence of phenomenal properties. This would not help the phenomenal realist, however. If phenomenal properties are the intrinsic natures of causally effective physical states, as panpsychists suppose, then they would have no distinctive effects, and we would still have to theorize about which physical states have phenomenal natures and what kind of phenomenal natures they have. Even if we discovered distinctive effects not predicted by physical theories, we could not establish that they were the product of irreducibly subjective states. Instead, we might treat the causes as theoretical posits, like other unobservables, defined by the causal role they play and therefore not irreducibly subjective.

If phenomenally conscious experiences have a special claim on our ethical concern, then this means that we can never be sure which experiences have this claim (except our own, and perhaps not even them if our introspective judgements are always memory-dependent). Nor would we have reason to trust our ethical instincts. The evolutionary processes that shaped these instincts could not have been directly sensitive to irreducibly subjective features, but only to physical correlates of them. And how could evolution

have been guided by the correct *metaphysical* theory of phenomenal–physical correlation? If ethical significance were written in a private phenomenal code to which we have no access, then we would never be able to regulate our interactions with other beings in a way that is sensitive to it.

It is worth stressing the oddity of this position — of locating a central source of ethical value in something to which our epistemic access is indirect and theoretically mediated. It would be comparable to locating the source of economic value in some inscrutable essence to which exchange value bore only a contingent and theoretically contentious relation. We couldn't build an economics on that basis, and it is hard to see how we could build an ethics on an analogous phenomenal one.

The illusionist perspective is very different. For the illusionist, the mental lives of others are not *irreducibly* subjective, and ethical status is not hostage to metaphysical theory. A creature's subjective life consists, not in the intrinsic quality of a private world, but in the dynamics of its engagement with the public one, and to know what it is “like” to be another creature, we shan't need to theorize about correlations between its physical and phenomenal states. We shall need to study its physical states very, very carefully, tracing the complex patterns of sensitivity and reactivity that constitute its subjective take on the world. And we shall need to devise frameworks for representing these patterns, creating multidimensional functional profiles for the experiences produced by different stimuli, which will provide an objective measure of how similar the creature's experiences are to our own along a variety of dimensions. (Again, I am using “functional” here in a broad sense; these profiles might include a lot of micro-level neural detail.) Such investigations might not tell us how to treat the creature, but they would give us all the information about consciousness relevant to our ethical decision making.

Distribution of Consciousness

A second implication concerns the distribution of consciousness in the natural world. Here phenomenal realists are committed to the existence of a sharp boundary. While they can allow that consciousness has different grades of richness and complexity in different creatures, they must make a sharp distinction

¹⁷ A reviewer objects that a cluster-based methodology of the sort advocated by Nicholas Shea and Tim Bayne might yield knowledge of the correlates of consciousness [42, 43]. Shea and Bayne point out that if consciousness is a natural kind, then there will be some underlying property that explains the clustering of the various perceptual capacities and reactive dispositions we treat as marks of consciousness, just as there may be an underlying pathology that explains the cluster of signs we associate with a disease. Identifying this property would then allow us to develop tests for consciousness, which could be used in cases where the usual marks are absent or inconclusive. I am not convinced that consciousness is a natural kind, but even if it is, this approach would not give us knowledge of *phenomenal* consciousness, understood as an irreducibly subjective property. For the natural kind property would be some objective neural or functional property, identified by its role in explaining certain psychological phenomena, and the metaphysical question of whether this property, and only this property, is accompanied by an irreducibly subjective what-it-is-likeness would remain as open as ever.

between those creatures that are conscious at all and those that aren't. A creature either has an irreducibly subjective aspect or it doesn't; there is no grey area. The show in the Cartesian Theatre may be minimal, but it's either there or not. Even panpsychists have to make such a distinction. Though they think that elementary particles have a tiny spark of phenomenal consciousness, they typically hold that these micro-consciousnesses combine to form rich macro-consciousnesses only in certain cases, combining in the brains of humans and other animals but not in tables and rocks. And there will be a sharp distinction between those creatures with a macro-consciousness and those without one.

If it is phenomenal consciousness that gives creatures intrinsic value, then this means that there will be a sharp ethical distinction between creatures that have some intrinsic value and ones that have none. Thus, phenomenal realists will have to divide organisms into the phenomenally conscious sheep, who matter intrinsically, and the phenomenally nonconscious goats, who don't. And they will have to do this from both phylogenetic and ontogenetic perspectives, identifying the points in evolution and development at which consciousness and intrinsic value emerge. These points will mark radical discontinuities in nature and radical differences in ethical status.

This is an unattractive position. Sharp discontinuities are rare in the biological world, and it is hard to think of any evolutionary or developmental point at which a principled line could be drawn between conscious and nonconscious beings. Moreover, the conviction that there must be such a line is potentially dangerous, opening space for religiously or politically motivated speculation about where it should be drawn.

Illusionism offers a different perspective. As illusionists see it, consciousness is not a binary feature — an inner light that is either on or off — but a set of complex, graded, multifaceted functional states. There is no sharp emergence of consciousness in either evolution or development, but rather a multidimensional continuity of perceptual and reactive complexity, starting with minimal forms, to which we would hesitate to apply the label “consciousness”, and ramifying into forms of consciousness much richer than our own. To navigate this complexity, we shall need to devise new taxonomic frameworks. Perhaps we could create functional profiles for the experiences

a creature undergoes in response to a standard range of stimuli and use them to define an overall *consciousness profile* for the creature, which could itself be located within a wider space of such profiles: we might call it *consciousness space*.¹⁸ Then, instead of asking whether or not a creature is conscious, we could ask where it is located in consciousness space, replacing a neat but intractable metaphysical question with a messy but tractable empirical one.

This will dictate a similar revision of our ethical practices. Instead of asking whether or not a creature has an ethical claim on us, we must ask what *kind* of ethical claim it has on us given its particular set of sensitivities and reactive dispositions. Instead of asking, “*Should* we care about this creature?”, we must ask, “*How* should we care about this creature?”.¹⁹

It might be objected here that illusionists are committed to a binary distinction of their own, between those creatures that undergo the introspective illusion of phenomenal consciousness and those that do not. And won't the former have a significantly different ethical status, in virtue of *seeming to have* an irreducibly subjective inner world?

This objection misunderstands the illusionist perspective — at least as I conceive of it. The illusion of phenomenal consciousness is not a substitute for phenomenal consciousness, which can do the same work in grounding ethical value. Illusionists think of conscious experiences as first-order functional states, whose valence is determined by the pattern of psychological reactions that constitute them. The ability to introspect these states does not confer some extra quality on them (a seeming what-it-is-likeness); it merely enables a raft of further psychological reactions with respect to them, including (if one is equipped with the right concepts) believing them to have an irreducibly subjective what-it-is-likeness. The illusion of phenomenal consciousness isn't a

¹⁸ Of course, one does not have to be an illusionist to believe in the possibility of developing consciousness profiles for different creatures, and detailed proposals along these lines have been made from a realist perspective [44, 45]. For the illusionist, however, such profiles will not be merely indications of a hidden subjective reality but maps of conscious experience itself.

¹⁹ For some highly interesting thoughts on the implications illusionism may have for questions about the distribution and ethical significance of consciousness, see Dung [38].

value-conferring inner glow; it is a representational process that induces us to believe in such an inner glow.

It is true that introspection adds new *objects* of awareness; creatures with introspective abilities are aware of their own mental states as well as of other aspects of themselves and their world. And we might call this a form of consciousness: introspective consciousness. But this is just another functional process, another layer of cognitive complexity, and it will come in different forms and grades (for dimensions along which possible introspective systems may vary, see [46]). Introspective consciousness will enrich the mental lives of the creatures that possess it, shifting their location in consciousness space, and this may affect how we treat them. While introspective capacities will not change the valence of a creature's first-order experiences, they may have effects that will alter the valence of its overall state. These effects will be multifarious. Reflecting on one's pain may cause additional suffering (for example, from worrying about one's health), but it may also offer sources of comfort and relief unavailable to less reflective creatures. All this should feed into our ethical decision making.

Ethical Progress

Consciousness is ethically important, but we are currently mystified about who possesses it. We know that humans are conscious, and we find it hard to doubt that other mammals are, but we are less sure about our more distant cousins, such as reptiles, and we draw a blank when it comes to beings very different from ourselves, such as insects, AIs, and, potentially, aliens. This is an unsatisfactory condition, and phenomenal realism offers little hope for improving it. It tells us that there are clear answers to questions about who and what is conscious but puts them tantalizingly out of our reach. And it encourages us to adopt a binary view of consciousness, which dismisses large swathes of living beings as insentient zombies, with no subjective life at all.

Illusionism, by contrast, offers a practical basis for deciding how to extend our ethical concern. It tells us that the facts about consciousness are empirically discoverable, and it encourages us to develop the tools needed to find and interpret them. The investigation won't be easy, and it will probably involve

distinguishing different types of consciousness, each with different aspects and degrees. It will be messy. That is all part of giving up the binary mindset of phenomenal realism.

Of course, empirical investigation on its own won't settle everything. It won't tell us how much weight to give to the suffering of other creatures and what trade-offs to make between their welfare and our own. No theory of consciousness would do that. But an illusionist approach promises to give our ethical decision making a much firmer basis, allowing us to establish the nature and valence of nonhuman experience through sensitive, painstaking scientific research, without the need to speculate about an inscrutable phenomenal code.

Conclusion

I opened this article with a quotation from the neuropsychologist Paul Broks, which can be seen as offering a vivid, even shocking, depiction of the illusionist perspective. Broks speaks of there being no self behind the face. What he means, I take it, is that there is no *immaterial consciousness* there — no “ethereal space”, no “essence”, distinct from the flesh and blood and brain. The only inner world is the one the neurosurgeon sees — the world of pulsating brain tissue. At first, this may seem a bleak, inhumane perspective, but as I have tried to show, it isn't. The impression of bleakness arises only if we look for consciousness in the wrong place, expecting to find it in some special entity or quality. It doesn't lie there, but in what the pulsating brain is *doing* — in the hugely complicated functions performed by the billions of massively interconnected neurons. And to understand the significance of this brain activity, we must take a wider perspective, or series of perspectives: organismic, environmental, interpersonal, and social. There may be no one there *in addition to* the brain, but the brain is the core of *someone* — the core of an amazingly complex conscious being, with a unique history, personality, and value. The conscious person isn't hidden in some inscrutable private realm but here in the public world with us — here for us to know and love and cherish.

Broks speaks of his realization as a liberation. And I think that is exactly right. The illusionist perspective liberates us. It liberates us from a conception of

ourselves as prisoners of private insubstantial worlds, which no one else can enter and from which we can never escape. It liberates us to really know our fellow creatures, human and nonhuman, and to apportion ethical concern more widely and more fairly within the wonderful natural world of which we are parts.

Acknowledgements The author thanks Daniel Dennett, Leonard Dung, François Kammerer, and two anonymous referees for their comments on an earlier version of this article.

Data Availability No datasets were generated or analysed in the course of preparing this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Broks, Paul. 2003. *Into the silent land: Travels in neuropsychology*. New York: Atlantic Monthly Press.
- Dennett, Daniel C. 1988. Quining qualia. In *Consciousness in modern science*, ed. Anthony J. Marcel and Edoardo Bisiach, 42–77. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198522379.003.0003>.
- Dennett, Daniel C. 1991. *Consciousness explained*. New York: Little, Brown and Co.
- Dennett, Daniel C. 2005. *Sweet dreams: Philosophical obstacles to a science of consciousness*. Cambridge, MA: MIT Press.
- Farrell, Brain A. 1950. Experience. *Mind* 59: 170–198.
- Feyerabend, Paul. 1963. Materialism and the mind-body problem. *The Review of Metaphysics* 17: 49–66.
- Frankish, Keith. 2016. Illusionism as a theory of consciousness. *Journal of Consciousness Studies* 23 (11–12): 11–39.
- Graziano, Michael S. A. 2013. *Consciousness and the social brain*. New York: Oxford University Press.
- Humphrey, Nicholas. 2011. *Soul dust: The magic of consciousness*. Princeton, NJ: Princeton University Press.
- Kammerer, François. 2021. The illusion of conscious experience. *Synthese* 198: 845–866. <https://doi.org/10.1007/s11229-018-02071-y>.
- Pereboom, Derk. 2011. *Consciousness and the prospects of physicalism*. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199764037.001.0001>.
- Rey, Georges. 1995. Towards a projectivist account of conscious experience. In *Conscious Experience*, ed. Thomas Metzinger, 123–142. Thorverton: Imprint Academic.
- Frankish, Keith. 2023. What is illusionism? *Klēsis Revue Philosophique* 55. <http://www.revue-klesis.org/pdf/klesis-55-04-keith-frankish-what-is-illusionism.pdf>. Accessed 16 May 2024.
- Nagel, Thomas. 1974. What is it like to be a bat? *The Philosophical Review* 83: 435–450. <https://doi.org/10.2307/2183914>.
- Tononi, Giulio, Melanie Boly, Marcello Massimini, and Christof Koch. 2016. Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience* 17: 450–461. <https://doi.org/10.1038/nrn.2016.44>.
- Goff, Philip. 2017. *Consciousness and fundamental reality*. New York: Oxford University Press.
- Chalmers, David J. 1996. *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press.
- Chalmers, David J. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2: 200–219. <https://doi.org/10.1093/acprof>.
- Carruthers, Peter. 1992. *The animals issue: Moral theory in practice*. Cambridge: Cambridge University Press.
- Searle, John R. 1999. *Mind, language and society: Philosophy in the real world*. London: Weidenfeld & Nicolson.
- Wilson, Catherine. 2018. Consciousness as a biological phenomenon: An alternative to panpsychism. *The Harvard Review of Philosophy* 25: 71–87. <https://doi.org/10.5840/harvardreview201891916>.
- Tononi, Giulio. 2008. Consciousness as integrated information: A provisional manifesto. *The Biological Bulletin* 215: 216–242. <https://doi.org/10.2307/25470707>.
- Strawson, Galen. 2008. Realistic monism: Why physicalism entails panpsychism. In *Real materialism: and other essays*, ed. Galen Strawson, 53–74. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199267422.003.0003>.
- Goff, Philip. 2016. The phenomenal bonding solution to the combination problem. In *Panpsychism: Contemporary perspectives*, ed. Godehard Bruntrup and Ludwig Jaskolla, 283–302. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199359943.003.0012>.
- Roelofs, Luke. 2019. *Combining minds: How to think about composite subjectivity*. New York: Oxford University Press.
- Shani, Itay, and Joachim Keppler. 2018. Beyond combination: How cosmic consciousness grounds ordinary experience. *Journal of the American Philosophical Association* 4: 390–410. <https://doi.org/10.1017/apa.2018.30>.
- Balog, Katalin. 2012. In defense of the phenomenal concept strategy. *Philosophy and Phenomenological Research* 84: 1–23. <https://doi.org/10.1111/j.1933-1592.2011.00541.x>.

28. Loar, Brian. 1990. Phenomenal states. In *Philosophical Perspectives, Vol. 4, Action Theory and Philosophy of Mind*, ed. James E. Tomberlin, 81–108. Atascadero, CA: Ridgeview. <https://doi.org/10.2307/2214188>.
29. Papineau, David. 2002. *Thinking about consciousness*. Oxford: Oxford University Press.
30. Carruthers, Peter, and Rocco Gennaro. 2023. Higher-order theories of consciousness. In *The Stanford Encyclopedia of Philosophy (Fall 2023 edition)*, ed. Edward N. Zalta and Uri Nodelman. <https://plato.stanford.edu/archives/fall2023/entries/consciousness-higher/>. Accessed 16 May 2024.
31. Frankish, Keith. 2020. Consciousness, attention, and response. *Cognitive Neuropsychology* 37: 202–205. <https://doi.org/10.1080/02643294.2020.1729111>.
32. Humphrey, Nicholas. 2006. *Seeing red: A study in consciousness*. Cambridge, MA: Harvard University Press.
33. Humphrey, Nicholas. 2023. *Sentience: The invention of consciousness*. Cambridge, MA: MIT Press.
34. Dennett, Daniel C. 2021. The user-illusion of consciousness. *Journal of Consciousness Studies* 28 (11–12): 167–177. <https://doi.org/10.53765/20512201.28.11.167>.
35. Strawson, Galen. 2019. A hundred years of consciousness: “A long training in absurdity.” *Estudios de Filosofía*: 9–43. <https://doi.org/10.17533/udea.ef.n59a02>.
36. Dennett, Daniel C. 2018. “Magic, illusions, and zombies”: An exchange. *The New York Review of Books* <https://www.nybooks.com/online/2018/04/03/magic-illusions-and-zombies-an-exchange/>. Accessed 16 May 2024.
37. Kammerer, François. 2020. The normative challenge for illusionist views of consciousness. *Ergo, an Open Access Journal of Philosophy* 6: 891–924. <https://doi.org/10.3998/ergo.12405314.0006.032>.
38. Dung, Leonard. 2022. Does illusionism imply skepticism of animal consciousness? *Synthese* 200: 238. <https://doi.org/10.1007/s11229-022-03710-1>.
39. Shabasson, Daniel. 2022. Illusionism about phenomenal consciousness: Explaining the illusion. *Review of Philosophy and Psychology* 13: 427–453. <https://doi.org/10.1007/s13164-021-00537-6>.
40. Balog, Katalin. The abolition of phenomena: A voyage among the zombies. *Kl̄esis Revue Philosophique* 55. <https://www.revue-klesis.org/pdf/klesis-55-07-katalin-balog-the-abolition-of-phenomena-a-voyage-among-the-zombies.pdf>. Accessed 16 May 2024.
41. Frankish, Keith. 2016. Not disillusioned: Reply to commentators. *Journal of Consciousness Studies* 23 (11–12): 256–289.
42. Bayne, Tim, and Nicholas Shea. 2020. Consciousness, concepts, and natural kinds. *Philosophical Topics* 48: 65–83. <https://doi.org/10.5840/philtopics20204814>.
43. Shea, Nicholas, and Tim Bayne. 2010. The vegetative state and the science of consciousness. *British Journal for the Philosophy of Science* 61: 459–484. <https://doi.org/10.1093/bjps/axp046>.
44. Birch, Jonathan, Alexandra K. Schnell, and Nicola S. Clayton. 2020. Dimensions of animal consciousness. *Trends in Cognitive Sciences* 24: 789–801. <https://doi.org/10.1016/j.tics.2020.07.007>.
45. Dung, Leonard, and Albert Newen. 2023. Profiles of animal consciousness: A species-sensitive, two-tier account to quality and distribution. *Cognition* 235: 105409. <https://doi.org/10.1016/j.cognition.2023.105409>.
46. Kammerer, François, and Keith Frankish. 2023. What forms could introspective systems take? A research programme. *Journal of Consciousness Studies* 30 (9–10): 13–48. <https://doi.org/10.53765/20512201.30.9.013>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.