

Keith Frankish

## *Illusionism as a Theory of Consciousness*

**Abstract:** *This article presents the case for an approach to consciousness that I call illusionism. This is the view that phenomenal consciousness, as usually conceived, is illusory. According to illusionists, our sense that it is like something to undergo conscious experiences is due to the fact that we systematically misrepresent them (or, on some versions, their objects) as having phenomenal properties. Thus, the task for a theory of consciousness is to explain our illusory representations of phenomenality, not phenomenality itself, and the hard problem is replaced by the illusion problem. Although it has had powerful defenders, illusionism remains a minority position, and it is often dismissed as failing to take consciousness seriously. This article seeks to rebut this accusation. It defines the illusionist programme, outlines its attractions, and defends it against some common objections. It concludes that illusionism is a coherent and attractive approach, which deserves serious consideration.*

So, if he's doing it by divine means, I can only tell him this: 'Mr Geller, you're doing it the hard way.' (James Randi, 1997, p. 174)

Theories of consciousness typically address the hard problem. They accept that phenomenal consciousness is real and aim to explain how it comes to exist. There is, however, another approach, which holds that phenomenal consciousness is an illusion and aims to explain why it *seems* to exist. We might call this *eliminativism* about phenomenal consciousness. The term is not ideal, however, suggesting as it does that belief in phenomenal consciousness is simply a theoretical error,

Correspondence:

Keith Frankish, The Open University, UK. Email: [k.frankish@gmail.com](mailto:k.frankish@gmail.com)

that rejection of phenomenal realism is part of a wider rejection of folk psychology, and that there is no role at all for talk of phenomenal properties — claims that are not essential to the approach. Another label is ‘irrealism’, but that too has unwanted connotations; illusions themselves are real and may have considerable power. I propose ‘illusionism’ as a more accurate and inclusive name, and I shall refer to the problem of explaining why experiences seem to have phenomenal properties as the *illusion problem*.<sup>1</sup>

Although it has powerful defenders — pre-eminently Daniel Dennett — illusionism remains a minority position, and it is often dismissed out of hand as failing to ‘take consciousness seriously’ (Chalmers, 1996). The aim of this article is to present the case for illusionism. It will not propose a detailed illusionist theory, but will seek to persuade the reader that the illusionist research programme is worth pursuing and that illusionists do take consciousness seriously — in some ways, more seriously than realists do.<sup>2</sup>

## 1. Introducing illusionism

This section introduces illusionism, conceived as a broad theoretical approach which might be developed in a variety of ways.

### 1.1. Three approaches to phenomenal consciousness

Suppose we encounter something that seems anomalous, in the sense of being radically inexplicable within our established scientific worldview. Psychokinesis is an example. We would have, broadly speaking, three options. First, we could accept that the phenomenon is real and explore the implications of its existence, proposing major revisions or extensions to our science, perhaps amounting to a paradigm shift. In the case of psychokinesis, we might posit previously unknown psychic forces and embark on a major revision of physics to accommodate them. Second, we could argue that, although the phenomenon is real,

<sup>1</sup> When I talk of phenomenal properties not being real or not existing, I mean that they are not *instantiated* in our world. This is compatible with the claim that they exist *qua* properties — a claim which illusionists need not deny.

<sup>2</sup> Defenders of illusionist positions (under various names) include Dennett (1988; 1991; 2005), Hall (2007), Humphrey (2011), Pereboom (2011), Rey (1992; 1995; 2007), and Tartaglia (2013). As Tartaglia notes, Place and Smart also denied the existence of phenomenal properties, which Place described as ‘mythological’ (Place, 1956, p. 49; Smart, 1959, p. 151).

it is not in fact anomalous and can be explained within current science. Thus, we would accept that people really can move things with their unaided minds but argue that this ability depends on known forces, such as electromagnetism. Third, we could argue that the phenomenon is illusory and set about investigating how the illusion is produced. Thus, we might argue that people who seem to have psychokinetic powers are employing some trick to make it seem as if they are mentally influencing objects.

The first two options are *realist* ones: we accept that there is a real phenomenon of the kind there appears to be and seek to explain it. Theorizing may involve some modest reconceptualization of the phenomenon, but the aim is to provide a theory that broadly vindicates our pre-theoretical conception of it. The third position is an *illusionist* one: we deny that the phenomenon is real and focus on explaining the appearance of it. The options also differ in explanatory strategy. The first is *radical*, involving major theoretical revision and innovation, whereas the second and third are *conservative*, involving only the application of existing theoretical resources.

Turn now to consciousness. Conscious experience has a subjective aspect; we say it is *like something* to see colours, hear sounds, smell odours, and so on. Such talk is widely construed to mean that conscious experiences have introspectable qualitative properties, or 'feels', which determine what it is like to undergo them. Various terms are used for these putative properties. I shall use 'phenomenal properties', and, for variation, 'phenomenal feels' and 'phenomenal character', and I shall say that experiences with such properties are *phenomenally conscious*. (I shall use the term 'experience' itself in a functional sense, for the mental states that are the direct output of sensory systems. In this sense it is not definitional that experiences are phenomenally conscious.) Now, phenomenal properties seem anomalous. They are sometimes characterized as simple, ineffable, intrinsic, private, and immediately apprehended, and many theorists argue that they are distinct from all physical properties, inaccessible to third-person science, and inexplicable in physical terms. (I use 'physical' in a broad sense for properties that are either identical with or realized in microphysical properties.) Again, there are three broad options.

First, there is radical realism, which treats phenomenal consciousness as real and inexplicable without radical theoretical innovation. In this camp I group dualists, neutral monists, mysterians, and those who appeal to new physics. Radical realists typically stress the anomalousness of phenomenal properties, their resistance to functional analysis,

and the contingency of their connection to their neural correlates. Second, there is conservative realism, which accepts the reality of phenomenal consciousness but seeks to explain it in physical terms, using the resources of contemporary cognitive science or modest extensions of it. Most physicalist theories fall within this camp, including the various forms of representational theory. Both radical and conservative realists accept that there is something real and genuinely qualitative picked out by talk of the phenomenal properties of experience, and they adopt this as their explanandum. That is, both address the hard problem.<sup>3</sup>

The third option is illusionism. This shares radical realism's emphasis on the anomalousness of phenomenal consciousness and conservative realism's rejection of radical theoretical innovation. It reconciles these commitments by treating phenomenal properties as illusory. Illusionists deny that experiences have phenomenal properties and focus on explaining why they seem to have them. They typically allow that we are introspectively aware of our sensory states but argue that this awareness is partial and distorted, leading us to misrepresent the states as having phenomenal properties. Of course, it is essential to this approach that the posited introspective representations are not themselves phenomenally conscious ones. It would be self-defeating to explain illusory phenomenal properties of experience in terms of real phenomenal properties of introspective states. Illusionists may hold that introspection issues directly in dispositions to make phenomenal judgments — judgments about the phenomenal character of particular experiences and about phenomenal consciousness in general. Or they may hold that introspection generates intermediate representations of sensory states, perhaps of a quasi-perceptual kind, which ground our phenomenal judgments. Whatever the details, they must explain the content of the relevant states in broadly functional terms, and the challenge is to provide an account that explains how real and vivid phenomenal consciousness seems. This is the illusion problem.

---

<sup>3</sup> Although all anti-physicalist theories are radical and all conservative theories physicalist, the radical/conservative distinction does not coincide with the anti-physicalist/physicalist one, since there may be radical physicalist theories.

### 1.2. *Illusionism strong and weak*

Illusionism makes a very strong claim: it claims that phenomenal consciousness is illusory; experiences do not really have qualitative, ‘what-it’s-like’ properties, whether physical or non-physical. This should be distinguished from a weaker view according to which some of the supposed *features* of phenomenal consciousness are illusory. Many conservative realists argue that phenomenal properties, though real, do not possess the problematic features sometimes ascribed to them, such as being ineffable, intrinsic, private, and infallibly known. Phenomenal feels, they argue, are physical properties which introspection misrepresents as ineffable, intrinsic, and so on. We might call this *weak illusionism*, in contrast to the strong form advocated here. (It might equally be called *weak realism*.)<sup>4</sup>

On the face of it, weak and strong illusionism are similar. Both hold that experiences have distinctive physical properties that are misrepresented by introspection. There is a crucial difference, however. Weak illusionism holds that these properties are, in some sense, genuinely *qualitative*: there really are phenomenal properties, though it is an illusion to think they are ineffable, intrinsic, and so on. Strong illusionism, by contrast, denies that the properties to which introspection is sensitive are qualitative: it is an illusion to think there are phenomenal properties at all.

We can highlight the difference by introducing the notion of a *quasi-phenomenal property*. A quasi-phenomenal property is a non-phenomenal, physical property (perhaps a complex, gerrymandered one) that introspection typically misrepresents as phenomenal. For example, quasi-phenomenal redness is the physical property that typically triggers introspective representations of phenomenal redness.<sup>5</sup> There is nothing phenomenal about such properties — nothing ‘feely’ or qualitative — and they present no special explanatory problem. Strong illusionists hold that the introspectable properties of experience are merely quasi-phenomenal ones. But weak

<sup>4</sup> For an example of a weak illusionist position, see Carruthers (2000, pp. 93–4, 182–91). For more examples, and discussion, see Frankish (2012), where phenomenal properties in this weakly illusionist sense are dubbed *diet qualia* — in contrast to *classic qualia*, or *qualia max*, which are genuinely ineffable, intrinsic, private, and so on. Compare also Levine’s distinction between *modest* and *bold* qualophilia (Levine, 2001, chapter 5).

<sup>5</sup> Extending the soft-drink metaphor, I have dubbed quasi-phenomenal properties *zero qualia* (Frankish, 2012).

illusionists cannot agree. If experiences have only quasi-phenomenal properties, then it would be misleading to say that phenomenal properties are real, just as it would be misleading to say that psychokinetic powers are real if all people can do is create the illusion of having them.

The moral is that if weak illusionism is not to collapse into strong illusionism, then it must employ a concept of phenomenality stronger than that of quasi-phenomenality. Indeed, one motive for advancing the strong illusionist position is to force conservative realists to face up to the challenge of articulating a concept of the phenomenal that is both stronger than that of quasi-phenomenality and weak enough to yield to conservative treatment. I doubt this is possible (see Frankish, 2012) and, if it is not, then radical realism and strong illusionism will be the only options. In what follows, ‘illusionism’ will always mean *strong illusionism*.

### 1.3. *Some analogies*

Illusionists offer various analogies to illustrate their view. Dennett compares consciousness to the user illusions created by the graphical interfaces through which we control our computers (Dennett, 1991, pp. 216–20, 309–14). The icons, pointers, files, and locations displayed on a computer screen correspond in only an abstract, metaphorical way to structures within the machine, but by manipulating them in intuitive ways we can control the machine effectively, without any deeper understanding of its workings. The items that populate our introspective world have a similar status, Dennett suggests. They are metaphorical representations of real neural events, which facilitate certain kinds of mental self-manipulation but yield no deep insight into the processes involved. (Dennett stresses the limits of the interface analogy. There is no internal display for the benefit of a conscious user; the illusion is a product of the limited access relations between multiple non-conscious subsystems and it manifests itself in our personal-level intuitions and judgments about our inner lives.)

Rey cites cases where stabilities in our reactions to the world induce us to project corresponding properties onto the world (Rey, 1995, pp. 137–9). For example, our stable personal concerns and reactions to others lead us to posit stable, persisting selves as their objects. Similarly, Rey suggests, our representations of our own and others’ experiences lead us to posit simple mental phenomena corresponding to them. Take pain, for example. We have a ‘weak’, functional concept

of pain, which includes links both to sensory representations of pain encoding information about intensity, apparent location, and so on, and to third-person representations of pain behaviour in others. Reflecting on our own and others' pains, we then develop a 'strong', qualitative concept of pain as the thing that is the immediate object of our pain experiences and the cause of pain behaviour in others.

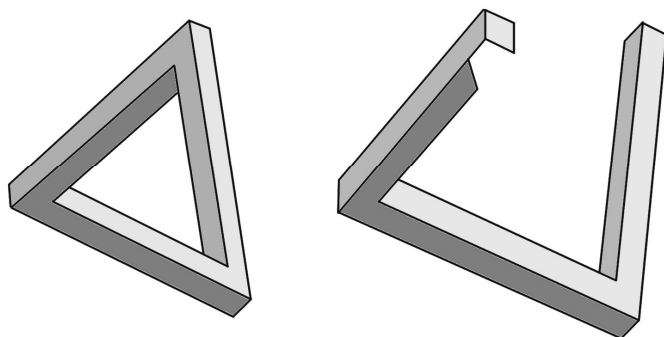


Figure 1. The Penrose triangle and the Gregundrum.

Humphrey compares sensations to impossible objects, such as the Penrose triangle, depicted on the left of Figure 1. Such an object cannot exist in three-dimensional space, but the illusion of it can be created by the object on the right, which Humphrey calls the *Gregundrum*, after its creator Richard Gregory. From most perspectives the Gregundrum appears an ungainly construction, but from just the right angle it looks like a solid Penrose triangle. Consciousness, Humphrey proposes, involves an analogous illusion. Our brains create an 'ipsundrum' — a neural state that appears relatively unremarkable from other perspectives but generates the illusion of phenomenality when viewed introspectively (Humphrey, 2011, chapter 2). Phenomenal consciousness is a '*fiction of the impossible*' (*ibid.*, p. 204) — a magic trick played by the brain on itself. (Talk of illusion should not be taken to indicate a defect in introspection; Humphrey argues that the illusion is highly adaptive; Humphrey, 2006; 2011.)

Pereboom draws a comparison with secondary qualities, such as colours (Pereboom, 2011, pp. 15–40). It is arguable that sensory perception represents colours as properties of external objects resembling the sensations they produce in us. Since objects lack such properties, sensory perception universally misrepresents objects in this respect.

Similarly, Pereboom suggests, introspection may universally misrepresent phenomenal properties as having qualitative natures they do not in fact have. (By 'phenomenal properties' he means the distinctive introspectable properties of conscious experiences, whatever they may be. Phenomenal properties in this sense may be merely quasi-phenomenal.) Pereboom calls this the *qualitative inaccuracy hypothesis*, and he argues that it is an open possibility. If it seems less credible than the parallel hypothesis about secondary qualities, Pereboom suggests, this is because we cannot check the accuracy of introspection, as we can that of perception, by adopting different vantage points, using measuring instruments, and so forth (*ibid.*, p. 23).

These analogies all illustrate the basic illusionist claim that introspection delivers a partial, distorted view of our experiences, misrepresenting complex physical features as simple phenomenal ones. Sensory states have complex chemical and biological properties, representational content, and cognitive, motivational, and emotional effects. We can introspectively recognize these states when they occur in us, but introspection doesn't represent all their detail. Rather, it bundles it all together, representing it as a simple, intrinsic phenomenal feel. Applying the magic metaphor, we might say that introspection sees the complex sleight-of-hand performed by our sensory systems as a simple magical *effect*. And, as with a conjuring trick, the illusion depends on what the audience does not see as much as what they do. In another analogy, Rey compares our introspective lives to the experience of a child in a dark cinema who takes the cartoon creatures on screen to be real (Rey, 1992, p. 308). The illusion depends on what the child doesn't see — on the fact that their visual system does not register individual frames as distinct images. Cinema is an artefact of the limitations of vision, and, illusionists may say, phenomenal consciousness is an artefact of the limitations of introspection.

The analogy with visual illusions also holds with respect to cognitive penetrability. Forming the theoretical belief that phenomenal properties are illusory does not change one's introspective representations, and one remains strongly disposed to make all usual phenomenal judgments (and perhaps does still make them at some level). As with perceptual illusions, this may indicate that the phenomenal illusion is an adaptive one, which has been hardwired into our psychology. (However, it may be possible to dispel the illusion



partially through indirect means, such as meditation and hypnotic suggestion; see, for example, Blackmore, 2011.)

The analogies also indicate some dimensions along which illusionist theories may differ. One concerns the sensory states that are the basis for the illusion. On most accounts, I assume, these will be representational states, probably modality-specific analogue representations encoding features of the stimulus, such as position in an abstract quality space, egocentric location, and intensity. Accounts will differ, however, on the details of their content, functional role, relation to attentional processes, and so on. Theories will also differ as to which properties of these states are responsible for the illusion of phenomenality (their quasi-phenomenal properties). Is introspection sensitive only to the content of sensory states, or are we also aware of properties of their neural vehicles? Do the reactions and associations evoked by our sensory states also contribute to the illusion of phenomenality?

Relatedly, there are questions about our introspective access to our sensory states. Do we have internal monitoring mechanisms that generate representations of sensory states, and if so what sort of representations do they produce? (Are they thoughts about sensory states or perceptions of their neural vehicles?) Are the introspective representations conscious or unconscious? (They are not phenomenally conscious, of course, but they could be conscious in the psychological sense of being globally available.) Are sensory states continually monitored or merely available to monitoring? Is the introspectability of sensory states a matter of internal access and influence rather than internal monitoring?<sup>6</sup> There are many options here and parallels with higher-order representational theories of consciousness, some of which might be reformulated as illusionist ones.

#### 1.4. *Outward-looking illusionism?*

I characterized illusionism as the view that phenomenal consciousness is an *introspective* illusion, reflecting the widely held view that phenomenal properties are properties of experience. This may be too restrictive, however. Some theorists hold that experience is *transparent*: when we attend to our experiences, we are aware only of properties of their objects. Thus, redness is experienced as a property

---

<sup>6</sup> On the varieties of introspection, see Prinz (2004).

of surfaces, pain as a property of parts of our bodies, and so on (e.g. Harman, 1990; Tye, 1995; 2000). This points to the possibility of an outward-looking illusionism, on which experience misrepresents distal stimuli as having phenomenal properties. Vision, for example, would represent objects as having illusory phenomenal colours as well as real physical colours (for a view of this kind, see Hall, 2007).

This view can be regarded as a variant of standard, inward-looking illusionism, differing principally on where the illusory phenomenal properties are represented as being located. And, like the inward-looking version, it may posit processes of internal monitoring. The illusion of phenomenality may involve a combination of introspection and projection, in which we both misrepresent features of experience as phenomenal and then re-represent these illusory properties as properties of the external world, mistaking complex physical properties of our sensory states for simple phenomenal properties of external objects (Humphrey, 2011, chapter 7). In what follows, I shall focus on the inward-looking form of illusionism, though most points will apply to both.

### *1.5. Illusionism and grand illusion*

Illusionism should be distinguished from the thesis that the visual world is a grand illusion (Noë, 2002). The latter holds that conscious visual experience is far less stable and detailed than we suppose, as is revealed by experiment and careful introspection. Illusionism, by contrast, is a thesis about conscious experience generally and concerns its nature, not its extent. One could hold that the visual world *is* stable and detailed while still claiming that it involves an illusion in the sense discussed here.

Nevertheless, evidence for the grand illusion view, such as the existence of change blindness, does lend support to illusionism. If we regularly overestimate the extent and stability of our conscious visual experience, then it is possible to be under a kind of illusion about one's own phenomenal consciousness. Moreover, as Dennett shows, phenomena such as change blindness undermine familiar intuitions about phenomenal properties, suggesting that our conception of them is incoherent and the properties themselves consequently illusory (e.g. Dennett, 2005, pp. 82–91).

### 1.6. *Illusionism and eliminativism*

Does illusionism entail eliminativism about consciousness? Is the illusionist claiming that we are mistaken in thinking we have conscious experiences? It depends on what we mean by ‘conscious experiences’. If we mean experiences with phenomenal properties, then illusionists do indeed deny that such things exist. But if we mean experiences of the kind that philosophers *characterize* as having phenomenal properties, then illusionists do not deny their existence. They simply offer a different account of their nature, characterizing them as having merely quasi-phenomenal properties. Similarly, illusionists deny the existence of phenomenal consciousness properly so-called, but do not deny the existence of a form of consciousness (perhaps distinct from other kinds, such as access consciousness) which consists in the possession of states with quasi-phenomenal properties and is commonly mischaracterized as phenomenal. Henceforth, I shall use ‘consciousness’ and ‘conscious experience’ without qualification in an inclusive sense to refer to states that might turn out to be either genuinely phenomenal or only quasi-phenomenal. In this sense realists and illusionists agree that consciousness exists.

Do illusionists then recommend eliminating talk of phenomenal properties and phenomenal consciousness? Not necessarily. We might reconceptualize phenomenal properties as quasi-phenomenal ones. Recall Pereboom’s analogy with secondary qualities. The discovery that colours are mind-dependent did not lead scientists to deny that objects are coloured. Rather, they reconceptualized colours as the properties that cause our colour sensations. Similarly, we might respond to the discovery that experiences lack phenomenal properties by reconceptualizing phenomenal properties as the properties that cause our representations of phenomenal feels — that is, quasi-phenomenal properties.<sup>7</sup> This could invite confusion, however, given how tightly the notion of phenomenality is bound up with dualist intuitions, and in scientific work it might be wiser to abandon talk of phenomenal properties and phenomenal consciousness altogether.

---

<sup>7</sup> Pereboom suggests that this might involve the unpacking of a conditional structure in phenomenal concepts (Pereboom, 2011, pp. 34–5). As he notes, a given phenomenal property might be reconceptualized either as the neural property that normally causes a representation of the relevant feel or as the higher-order property of being a neural state that could cause a representation of it (*ibid.*).

In everyday life, however, we would surely continue to talk of the feel or quality of experience in the traditional, substantive sense. As subjects of experience, our interest is in how things seem to us introspectively — the illusion itself, not the mechanisms that cause it. Such talk may fail to pick out real properties, but it is not empty or pointless. Consider another analogy. Having watched a performance of *King Lear*, Lucy remarks, ‘Lear’s anguish in the final scenes was heart-breaking’. What is she talking about? There was (we may suppose) no anguish on stage at all, only the artful illusion of it. And it would be implausible to construe Lucy as referring to the *cause* of this illusion — the actor’s words and gestures (quasi-anguish, as it were). The words and gestures were not themselves heart-breaking. The answer, of course, is that Lucy is referring to a fictional agony, entering into the world of the play and responding to the emotions of the characters as if they were real. (And in doing so, we might add, she is not making an error but appreciating the very point of the performance.) Everyday talk about the quality of experience should, I suggest, be construed similarly. Of course, most people do not regard their phenomenology as illusory; they are like naïve theatregoers who take the action on stage for real. But if illusionists are right, then cognitive scientists should treat phenomenological reports as fictions — albeit ones that provide clues as to what is actually occurring in the brain.<sup>8</sup>

### *1.7. Zombies and what it is like*

Are illusionists claiming that we are (phenomenal) zombies? If the only thing zombies lack is phenomenal consciousness properly so called, then illusionists must say that, in this technical sense, we are zombies. However, zombies are presented as creatures very different from ourselves — ones with no inner life, whose experience is completely blindsighted. As Chalmers puts it, ‘There is nothing it is like to be a zombie... all is dark inside’ (Chalmers, 1996, pp. 95–6). And illusionists will not agree that this is a good description of us. Rather,

---

<sup>8</sup> Compare Dennett’s story of the forest god Feenoman (Dennett, 1991, chapter 4). Local tribespeople believe Feenoman is real, but visiting anthropologists treat him as an intentional object, defined by the locals’ beliefs, and remain neutral on the question of what lies behind the myth. Dennett recommends that we treat first-person phenomenological reports in the same way, as data for third-person theorizing (‘heterophenomenology’).

they will deny the equivalence between having an inner life and having phenomenal consciousness. Having the kind of inner life we have, they will say, consists in having a form of introspective self-awareness that creates the illusion of a rich phenomenology.

But aren't phenomenal properties precisely what makes experience *like something*? That is certainly a common way of construing what-it's-like talk, but there is another way. Illusionists can say that one's experiences are like something if one is aware of them in a functional sense, courtesy of introspective representational mechanisms. Indeed, this is a plausible reading of the phrase; experiences are like something for a creature, just as external objects are like something for it, if it mentally represents them to itself. Illusionists agree that experiences are like something in this sense, though they add that the representations are non-veridical, misrepresenting experiences as having phenomenal properties (what-it's-like-ness in the first sense). And in this second sense there *is* something it is like to be a zombie, since zombies have introspective mechanisms functionally identical to our own. When we imagine zombies as being different from us, we are — illegitimately — imagining creatures with different introspective capacities.

It may be objected that we can imagine a creature representing itself as having phenomenal properties while still lacking an inner life. Zombies believe they are phenomenally conscious (in some sense at least; arguably, they lack full-blown phenomenal concepts; Chalmers, 1996; 2003). But — it may be said — this does not give them an inner life like ours. I am not sure this is obvious. Consider the grand illusion view again. Our sense that our visual field is uniformly rich and detailed may be a sort of cognitive illusion, reflecting expectations and assumptions about the information that vision provides, and our sense of having a rich phenomenology might be a similar cognitive illusion. But in any case the illusionist need not claim that the illusion depends solely on the possession of certain propositional attitudes. Rather, they may say, it depends on a complex array of introspectable sensory states, which trigger a host of cognitive, motivational, and affective reactions. If we knew everything about these states, their effects, and our introspective access to them, then, illusionists say, we could not clearly imagine a creature possessing them without having an inner life like ours.

Of course, it is easy to say that. Illusionists need to explain how it can be true. That is, they need to solve the illusion problem. But it

would be begging the question against illusionism to assume that it cannot be done.

## 2. Motivating illusionism

This section motivates illusionism, sketching its advantages over radical realism and conservative realism and then adding some positive arguments in its favour. It does not aim to present a watertight case for illusionism but simply to show that the view has strong attractions.

### 2.1. *Against radical realism*

I take it there is a presumption in favour of conservatism in science: we should not make radical theoretical moves if modest ones will do. Of course, when it comes to consciousness many are confident that modest moves *won't* do, but that is what conservative theorists deny. The principle of conservatism should apply with special force, I suggest, when the pressure for radical innovation comes from a parochial, anthropocentric source, such as introspection. Introspection delivers a view of ourselves that is peculiarly vivid and compelling and that seems radically at odds with that of the physical sciences. It *might* give us access to an aspect of reality inaccessible to third-person science. (Though even if it did, it is hard to see how we could develop a science of that aspect.<sup>9</sup>) But it might merely give us an unusual perspective on the same reality — a perspective that is partial and distorted and deceives us into thinking that our experiences are resistant to conservative explanation.

In addition, a conservative approach is much better placed to account for the *psychological significance* of consciousness. By the psychological significance of a mental event, I mean its cumulative cognitive, motivational, emotional, and other psychological effects across various contexts. The common-sense view is that the way our experiences feel has huge psychological significance. Sensations entice us, guide us, move us, warn us, and the memory and anticipation of them are powerful motivators. Not only this, they hugely *enrich* life. As Humphrey stresses, we relish sensation for its own sake, and this relish shapes our behaviour in profound ways

---

<sup>9</sup> For the case against first-person science, see Dennett (1991, chapter 4; 2003; 2005, chapter 6; 2007).

(Humphrey, 2011). But this assumes that experiences affect us in virtue of how they feel. And it is hard for radical theorists to vindicate this assumption. Non-physical properties can have no effects in a world that is closed under causation, as ours appears to be, and the mind sciences show no independent need to refer to exotic physical processes, such as quantum-mechanical ones. The threat of epiphenomenalism hangs over radical theories. Some radical theorists respond by arguing that phenomenal properties are intrinsic to basic physical entities and thus intimately involved in physical causal processes (e.g. Strawson, 2006). However, even if this proposal does dispel the threat (which is doubtful; Howell, 2015), it involves huge profligacy with phenomenal properties and preserves the potency of consciousness only at the cost of making all physical causation phenomenal.

## 2.2. *Against conservative realism*

Conservative realism promises to capture the common-sense view of consciousness, accepting the reality of phenomenal properties but identifying them with causally potent, physical properties. However, it is an unstable position, continually on the verge of collapsing into illusionism.

The central problem, of course, is that phenomenal properties seem too weird to yield to physical explanation. They resist functional analysis and float free of whatever physical mechanisms are posited to explain them. (In practice this becomes almost definitional of phenomenal consciousness; any physicalist theory can be rejected as missing out the essential qualitative element.) The arguments are well-known, and I shall not repeat them here.

Many physicalists respond by arguing that our anti-physicalist intuitions arise from the way we conceptualize phenomenal properties in introspection — a tactic known as the *phenomenal concept strategy* (e.g. Hill, 1997; Hill and McLaughlin, 1999; Loar, 1990; Papineau, 2002). The idea is that phenomenal concepts have an especially intimate link to their referents and lack *a priori* connections to physical concepts. (They are typically characterized as either demonstrative, recognitional, or quotational.) This intimacy and isolation, it is argued, give rise to anti-physicalist intuitions, even though phenomenal properties are physical ones. It is doubtful, however, that this really relieves the pressure on conservative realism. For the concepts must still be *phenomenal* ones (Tartaglia, 2013, p. 828). If they are recognitional concepts, for example, they must be recognitional concepts

for the *feel* of experiences. The concept of a mere introspectable *something*, which might or might not be qualitative, is not a genuine phenomenal concept, and if we conceptualized the properties of experience in that way, we would not feel any resistance to thinking of them as physical (a bare something might as easily be physical as non-physical). But if phenomenal concepts refer to feels, then the challenge to conservative realists remains. They must either explain how these feels can be physical or accept that phenomenal concepts misrepresent experience, as illusionists claim.

Looking at proposed reductive explanations themselves, the pressures towards illusionism become even clearer. As noted earlier, most physicalists adopt a weakly illusionist view, denying that phenomenal properties are private, intrinsic, and ineffable and employing the phenomenal concept strategy to explain why they seem so. However, they insist that phenomenal properties are nonetheless real and genuinely qualitative. I have already suggested that this position is problematic. If it is not to collapse into illusionism, then it must employ a notion of phenomenality that is stronger than that of quasi-phenomenality. Phenomenal properties must not merely cause representations of phenomenality but have some genuinely ‘feely’ aspect to them. And it is unclear what this could be. What phenomenal residue is left, once features such as privacy, intrinsicity, and ineffability have been stripped away (Frankish, 2012)?

In practice, reductive explanations of phenomenality tend to take a covertly illusionist form. They typically identify phenomenal character with some functional property of experience such as possession of a certain kind of representational content or availability to higher-order representation. But in so far as these identifications are plausible, it is, I suggest, because subjects whose experiences had this functional property would be disposed to judge that their experiences had a qualitative dimension, rather than because their experiences really would have such a dimension. In the case of higher-order perception theory, for example, it may be true that perceptual awareness of the physical vehicles of experience would create the sense that experiences have an intrinsic quality. But this is an explanation of quasi-phenomenal properties, not phenomenal properties. There is a conflation of phenomenality with the representation of phenomenality, and thus of realism with illusionism.

Of course, these objections assume that we are seeking an *explanation* of consciousness. Physicalists can resist illusionist pressures if they are content to accept the existence of an *explanatory*



*gap* between phenomenal properties and their neural substrates (e.g. Levine, 2001). Others, however, may prefer an explicable illusion to an inexplicable reality.

It may be objected that illusionism discards one of the major advantages of conservatism, namely that it gives phenomenal properties a causal role. If phenomenal properties are illusory, then they have no causal role after all. Illusionists can reply that they do not deny that phenomenal concepts track causally effective properties; they merely deny the common-sense view of the nature of these properties — that they are qualitative. Or, perhaps more persuasively, they can say that phenomenal properties *are* causally potent, considered as intentional objects. They move us in the same way that ideas, stories, theories, and memes do, by figuring as the objects of our intentional states. In talking of the power of sensation we are talking of the power of certain representational contents.

### 2.3. *For illusionism*

The case for illusionism can also be made in a positive way, appealing to explanatory considerations. If phenomenal consciousness is conceived as non-physical, then, as Chalmers notes, there is a simple argument for its being illusory (Chalmers, 1996, pp. 186–7: Chalmers does not endorse the argument, of course, though he acknowledges its force). If people's claims and beliefs about something (God, say, or UFOs) can be fully explained as arising from causes having no connection with the thing itself, then this is a reason for discounting them and regarding the thing as illusory. But it is widely accepted, even by anti-physicalists, that we do not need to appeal to non-physical properties in order to explain our behaviour and the mental processes that cause it, including our assertions and beliefs about our own conscious experiences. Phenomenal zombies would make the same assertions we do about their conscious experiences and about consciousness in general, and with the same conviction, and they would have beliefs on those matters with the same causal and explanatory roles as ours (though, arguably, with different contents). Given this, our claims and beliefs about consciousness afford no evidence for the truth of phenomenal realism, and it is reasonable to regard them as mistaken.

A second argument for illusionism does not depend on the assumption of anti-physicalism. In general, apparent anomalousness is evidence for illusion. If a property resists explanation in physical

terms or is detectable only from a certain perspective, then the simplest explanation is that it is illusory. In this light, considerations usually cited in support of a radical approach to consciousness, such as the existence of an explanatory gap, the conceivability of zombies, and the perspectival nature of phenomenal knowledge, afford equal or greater support for illusionism. Given the force of these considerations, if there is even a remote possibility that we are mistaken about the existence of phenomenal consciousness, then there is a strong abductive inference to the conclusion that we are in fact mistaken about it. And there is reason to think that we could be mistaken about it. For our awareness of phenomenal properties would have to be mediated in some way. If the mind is a representational system, then properties must be mentally represented in order to have cognitive, affective, or motivational significance, and phenomenal properties are no exception, regardless of whether they are physical or non-physical. A creature that lacked introspective representations of its phenomenal properties — we might call it a *representational zombie* — would have no cognitive access to its phenomenal properties and would be unable to form beliefs about them, reflect on them, report them, remember them, respond emotionally to them, or act upon them. Its experiences would not be *like anything*, in the second of the senses distinguished earlier.<sup>10</sup> But we have no introspective way of checking the accuracy of our introspective representations, and so cannot rule out the possibility that they are non-veridical. (Indeed, in so far as we can check, through external inspection of our brain states, they appear to be non-veridical; the properties represented do not show up from other perspectives.) For all we know, then, phenomenality might be illusory; and, given its anomalousness, we can abductively infer that it is.

Illusionism has other explanatory advantages too. One is that it permits us to acknowledge both the wonder of phenomenal consciousness and its potency. This is something realists find hard to do. Stressing the magical, non-physical character of phenomenal properties usually means denying them a causal role, while treating them as physical causes means denying that they are as magical as they seem.

---

<sup>10</sup> Compare Rey: 'Postulating qualia properties, whether in the brain or in some special realm, will be of no help unless we have an account of how those properties are assimilated into a person's cognitive life; and it's hard to see how they could be assimilated without being *represented*' (Rey, 2007, pp. 129–30).

But if phenomenal properties are intentional objects, a sort of mental fiction, then we need no longer be embarrassed by them. We can acknowledge how magical and unearthly they are *and* how powerfully they affect us, as intentional objects. In this sense, illusionists may claim to take consciousness more seriously than realists do.

Illusionism also offers an attractive perspective on the function of consciousness. If consciousness has the powerful behavioural influence it seems to have, then we should be able to explain it as an adaptive feature. Again, realists find this hard to do. If consciousness is a matter of pure feel, then it is unclear what function it could perform, and many realists, both radical and conservative, see it as little more than a side effect of perceptual processes. But if consciousness involves an illusion, then new possibilities open. Maybe its function is precisely to give us the impression that we have a magical, non-physical inner life. Humphrey has made a powerful case for such a view (1992; 2006; 2011). He proposes that sensations occur when internalized evaluative responses to stimuli ('sentitions') interact with incoming sensory signals to create complex feedback loops, which, when internally monitored, seem to possess otherworldly, phenomenal properties. This internal 'magic show', Humphrey argues, powerfully affects the creatures that possess it, giving them a new interest in their existence, inducing them to engage more deeply with their environment (onto which they project phenomenal properties), and creating a sense of self, and, in humans, belief in an ego or immaterial soul. These developments, Humphrey argues, were strongly fitness-enhancing, and the magic show has been sculpted by natural selection to promote them. (This is possible since, on Humphrey's view, the mechanisms of sensation are separate from those of perception and can respond to different evolutionary pressures.) Whether or not this account is right (and it has many attractions), it is an excellent illustration of how evolutionary theorizing about consciousness can flourish, once freed from the metaphysical preoccupations of realism.

### 3. Defending illusionism

This section responds to some common objections to illusionism. It argues that they serve primarily to highlight the commitments of the illusionist approach and that illusionists can accommodate weakened versions of the intuitions on which they draw.

### 3.1. *Denying the data*

The most basic objection to illusionism is that it denies the data. To be sure, if all that needed to be explained were the detectable marks of phenomenal consciousness — the related judgments, reports, reactions, dispositions, and so on — then it would be more economical to adopt an illusionist view. But — the objection goes — that is not all that needs to be explained; phenomenal consciousness is itself a datum (Chalmers, 1996, p. 188). Phenomenal properties are not theoretical posits introduced to explain other data, but are themselves core data.

There is a sense in which illusionists can agree. It is a datum that phenomenal properties exist as intentional objects; our introspective reports define a notional introspective world which is as we take it to be. But illusionists do, of course, deny that phenomenal properties exist in the real world, as properties of brain states. We are strongly disposed to think that their existence is an introspective datum, but all observation statements, including ones about our own minds, are open to revision in the light of theory. Our introspective reports are data for a science of consciousness, but they require interpretation and evaluation, and the best explanation for them may be one that denies their reliability (Dennett, 2003; 2007). And, as we have seen, there are strong theoretical reasons to doubt the reliability of our first-person reports about phenomenal consciousness.

If realists are to maintain that phenomenal consciousness is a datum, then they must say that we have a special kind of epistemic access to it, which excludes any possibility of error. And since no causal process could provide such certainty, they must say that this access is not causally mediated. This is indeed what some realists propose. Chalmers holds that we are *directly acquainted* with phenomenal properties (Chalmers, 1996, pp. 192–200). He describes acquaintance as ‘a basic sort of epistemic relation between a subject and a property’ and says that ‘whenever a subject has a phenomenal property, the subject is acquainted with that phenomenal property’ (2003, p. 250). Experience is in this sense intrinsically epistemic (1996, p. 196).

This view protects the status of phenomenal consciousness as a datum but does so at a high cost. First, acquaintance can have no psychological significance. In order to talk or think about our phenomenal properties, we need to form mental representations of them, and since representational processes are potentially fallible, the certainty conferred by acquaintance could never be communicated,

either to others or even to ourselves, considered as cognitive systems. The price of making consciousness a datum is that the datum is psychologically inert. Second, acquaintance theory assumes that the reactions and associations a sensory episode evokes do not affect its feel, since we are not directly acquainted with them or their effects. Yet there is reason to think that our reactions and associations do shape our sense of what our experiences are like (see Dennett, 1988; 1991, chapter 12; 2005, chapter 4). (It might be replied that these factors influence our *judgments* about the feel of our experiences, not the feels themselves, but this would open a systematic gap between what our experiences are like and what we think they are like — which is, at the least, counter-intuitive.)

Acquaintance theory also comes with heavy metaphysical baggage. It is hard to see how physical properties could directly reveal themselves to us, so the theory plausibly assumes an anti-physicalist view of phenomenal consciousness. Moreover, it may require an anti-physicalist view of the experiencing subject too. If subjects are complex physical organisms, how can they become directly acquainted with phenomenal properties? When cognitive scientists talk of information being *available to the subject*, they mean that it is globally broadcast, available for the flexible control of thought and action, and so on. But events need to be represented in order to be available to the subject in this sense. Talk of acquaintance supposes a non-psychological subject, which exists prior to representational processes, as opposed to being partially constituted by them.

This brings us back to talk of it being *like something* to be us. As noted earlier, such talk may mean simply that we have an introspective awareness of our experiences, generated by representational mechanisms. We might call this *introspective subjectivity*. Illusionists agree that we have introspective subjectivity, though they hold that it is radically misleading. But ‘like something’ talk can be understood in a stronger sense, as indicating that we possess a subjective dimension that is not the product of introspective mechanisms but arises simply from our being the things we are. Call this *intrinsic subjectivity*. When theorists talk of our being directly acquainted with phenomenal properties, it is intrinsic subjectivity they have in mind; the properties, and our awareness of them, are simply correlates of our physical constitution. Plausibly, then, taking phenomenal consciousness as a datum involves positing intrinsic subjectivity.

Intrinsic subjectivity is, however, deeply mysterious. It is a shadowy companion of physical systems, and we could imagine *any* object

possessing it, as panpsychists do. (It might be proposed that only beings with a certain physical structure possess intrinsic subjectivity — perhaps only those that implement an information processing system. But this structure does not *explain* their intrinsic subjectivity, and a creature's reports of what its experiences are like will be the product of introspective mechanisms and will thus manifest introspective subjectivity only.) This does not show that the notion of intrinsic subjectivity is incoherent, but it is, I think, a good reason to explore the idea that it is a fiction created by introspective subjectivity.

### 3.2. *No appearance–reality gap*

Another common objection to illusionism is that in the case of qualitative states there is no gap between illusion and reality. Something can look like a Penrose triangle without being a Penrose triangle, but an experience that seems to have a greenish phenomenal character really does have a greenish phenomenal character.<sup>11</sup> As Searle puts it, '*where consciousness is concerned the existence of the appearance is the reality*'. If it seems to me exactly as if I am having conscious experiences, then I am having conscious experiences' (Searle, 1997, p. 112, italics in original).<sup>12</sup>

This is often presented as a crushing objection to illusionism, but it is far from compelling. It turns on what we mean by *seeming to have* a greenish experience. If we mean having an introspective experience with the same phenomenal feel as a greenish experience, then, trivially, there is no distinction between seeming and reality. But of course that is not what illusionists mean. They mean introspectively representing oneself as having a greenish experience, and one can do this without having a greenish experience. The objector may reply that, in order to create the illusion of a greenish experience, the introspective representation would have to employ a greenish mode of presentation, which would itself have an introspectable greenish feel. However, illusionists will simply deny this, arguing that the content of introspective representations is determined by non-phenomenal,

<sup>11</sup> I follow Levine's practice of using 'greenish' for the (putative) feel associated with perception of a green object (Levine, 2001).

<sup>12</sup> Compare Kripke: 'in the case of mental phenomena there is no "appearance" beyond the mental phenomenon itself' (Kripke, 1980, p. 154).

causal or functional factors.<sup>13</sup> The objector may say that there is a big difference between merely representing oneself as having a greenish experience in such a way and actually having a greenish experience, but that is just the point at issue. The illusionist claims that when we think we are having a greenish experience we are in fact merely misrepresenting ourselves as having one. That claim may be false, but the no-gap objection does not add anything to the case against it. Of course, this requires some account of the content of the representations involved, and providing this will be a major challenge for the illusionist. But it is an independent requirement, and the no-gap objection does not make it harder to meet.

Another version of the no-gap objection might go as follows.<sup>14</sup> It is incoherent to doubt that experiences are as they seem, since experience reports are already reports of how things seem. I may come to doubt my initial claim that there is a green patch in front of me and retreat to the more cautious claim that there seems to be a green patch, but I cannot coherently retreat from that claim to the claim that there *seems to seem* to be a green patch. The first claim expresses all the epistemic caution that is necessary or possible. There is something right about this. We have no everyday procedure for correcting sincere and attentive experience reports, and we treat them as authoritative. But it does not follow that this authority is epistemic. Being cautious about the external world does not make one authoritative about the internal one, and seeming to see a green patch isn't the same as infallibly introspecting a greenish phenomenal feel. Rather, as Dennett suggests, the authority might be more like that which a storyteller has over their fictions (Dennett, 1991, p. 81).<sup>15</sup>

---

<sup>13</sup> Alternatively, illusionists might concede that introspection employs modes of presentation that appear to have phenomenal feels, but argue that this too is an illusion — that introspection misrepresents the modes of presentation as having phenomenal properties they lack. For defence of this option and an argument that it does not generate an infinite regress, see Pereboom (2011, pp. 27–8).

<sup>14</sup> This version was suggested by remarks of Martine Nida-Rümelin, though she might not endorse my presentation of it.

<sup>15</sup> It might be argued that phenomenal properties cannot be illusory, since they serve as sense-data, and it is only when sense-data are interpreted that illusion can arise (Wright, 2008). This is unpersuasive, however, even granting sense-data theory. Introspective representations of phenomenal properties might serve as data in the construction of representations of external reality while themselves misrepresenting internal, neurophysiological reality. (I am grateful to an anonymous referee for raising this objection.)

In a strong form, then, the claim that there is no appearance–reality gap for phenomenal properties is not compelling. A weaker version of the claim is, however, both plausible and compatible with illusionism. From the perspective of a representational theory of mind, the difference between seeming to be aware of a certain phenomenal feel and actually being aware of it is that between having a non-veridical introspective representation of the feel and having a veridical one, and, subjectively, this is no difference at all. In this sense, illusionists can agree that there is no appearance–reality gap for consciousness.

### 3.3. *Who is the audience?*

An illusion presupposes an audience. Who is the audience for the illusion of phenomenal consciousness? Illusionists will join Dennett in dismissing the idea that there is an inner arena (a ‘Cartesian theatre’) where perceptual information is assembled and a phenomenal show presented for an appreciation by an internal observer (Dennett, 1991). But aren’t they committed to reinstating a Cartesian theatre as an arena where the *illusion* of phenomenality is presented?

Illusionists may be committed (as many theorists are) to the existence of an inner *representer* of some kind: a system, or set of processes, which generates introspective representations of sensory states. But this need not amount to an observer, still less a conscious one. If we do not need an inner observer to appreciate perceptual representations, why should we need one to appreciate introspective ones? As Dennett argues, once the brain has made a discrimination, there is no need for another brain system to remake it, and all the work of appreciation and reaction can be (and ultimately must be) distributed among numerous unintelligent subsystems (*ibid.*). Similarly, once an introspective representation has been generated, the work of reacting to it — of being impressed by the illusion — can, and must, be distributed across such subsystems. There need be no unified audience for the illusion smaller than the organism as a whole (or at least its central nervous system).

That said, illusionists may posit something like an inner display. Recall Humphrey’s proposal that internal feedback loops have been shaped by evolution to create a life-enhancing internal magic show (Humphrey, 2011). Such a show is, however, different from the one in the Cartesian theatre. First, it is not a phenomenal show, though it is represented as one. Second, it is not a redundant re-presentation of information already encoded in the system. The feedback loops are



new features, continuously generated, which need to be monitored and represented in order to have psychological effects. Third, the detector system need do no more than generate representations; again, all the work of appreciating and reacting to the show can be parcelled out to subsystems. Finally, (though Humphrey might not agree) the show need not be a single, integrated one, generating a definitive stream of introspective representations. Instead, there might be numerous micro-presentations, yielding multiple drafts of sensation (an introspective counterpart of the multiple drafts model of perceptual processing Dennett proposes; Dennett, 1991). Extending the theatrical metaphor, there might be a host of fringe events around the town rather than an official show in a central auditorium.

### *3.4. Representing phenomenality*

Another objection centres on the representation of phenomenality. If there are no phenomenal properties, how do we represent them? How do we acquire phenomenal concepts, and how do these concepts capture the richness of phenomenality? These are central questions for illusionists, and answering them would go a long way towards solving the illusion problem. Here I shall merely make some preliminary remarks and indicate some lines open to the illusionist.

The task of constructing a theory of content for phenomenal concepts is a difficult one, but it is not obvious that it is significantly more difficult for those who hold that these concepts lack referents. Levine questions whether we can explain the richness and determinacy of our phenomenal representations without reference to actual phenomenal properties (Levine, 2001, pp. 146–7). When we think about what an experience is like, he suggests, the phenomenal property itself is included in the thought and serves as its own mode of presentation (*ibid.*, p. 8). The idea that phenomenal concepts quote or incorporate tokens of their referents has been proposed by several theorists (e.g. Chalmers, 2003; Papineau, 2002, pp. 116–25). However, its explanatory power is questionable. Why should incorporating a phenomenal feel into a representational vehicle make the vehicle represent the feel, let alone in a rich and determinate way? (Incorporating iron filings into it wouldn't make it represent iron.) As Rey stresses, some mechanism would be needed to read off features of the incorporated property and represent them to the rest of the system (Rey, 2007, pp. 128–9). But then a secondary, non-quotational account of phenomenal

representation would be needed, to which the illusionist could appeal directly.

It is true that illusionism does not sit well with strong externalist views, on which the content of a representation is constituted by causal connections to its referent. Illusionists might argue that phenomenal concepts are compounded from more primitive ones that do refer, or that they have counterfactual causal connections to uninstantiated phenomenal properties. However, there are reasons for finding neither of these options attractive, either for phenomenal concepts or for non-referring concepts generally (Rey, 2005). A better option may be to adopt some form of functional-role semantics for phenomenal concepts, on which their content is fixed by their role in mental processing, including their connections to other concepts, to non-conceptual sensory and introspective representations (their own content determined causally or functionally), and to associations, behavioural dispositions, and so on. (If these functional roles are narrow, ‘in the head’ ones, the content of our phenomenal representations will be independent of environmental factors — but that is not implausible; see Rey 1998.)

Another possibility is that phenomenal concepts are hybrid ones. Suppose we have a general theoretical concept of a phenomenal property — roughly, that of a simple, intrinsic, immediately known, introspectable property of experience. This concept might be innate, the product of individual theorizing, or culturally acquired. Suppose, too, that we have capacities to introspectively recognize different types of sensory states when they occur, and associated recognitional concepts for the states identified. Then phenomenal concepts might be hybrid ones combining the general theoretical concept with specific recognitional ones. For example, the concept of a certain shade of phenomenal red might be that of *this kind of phenomenal property*, where ‘this kind’ refers to the kind picked out by the recognitional capacity exercised while having an experience of the relevant type. Of course, if illusionism is true, that capacity does not pick out a phenomenal property; it picks out a complex physical one. So the hybrid concept fails to refer. (Compare ‘that kind of ectoplasm’ said by a credulous spectator at a séance.) Indeed, the theoretical concept may inform our introspective awareness, so that we mistakenly introspect sensory states *as* phenomenal, just as we might misperceive a flat hologram *as* a three-dimensional object (perhaps even an impossible one, such as a Penrose triangle). A hybrid theory like this may be able to account for many of our intuitions about phenomenal

consciousness, rendering illusionism more palatable. If introspection employs recognitional concepts, it may present its objects as being simple, ineffable, and immediately known, but if it is also theoretically informed, it may at the same time radically misrepresent them.

#### 4. Facing up to the illusion problem

Illusionism replaces the hard problem with the illusion problem — the problem of explaining how the illusion of phenomenality arises and why it is so powerful. This problem is not easy but not impossibly hard either. The method is to form hypotheses about the underlying cognitive mechanisms and their bases in neurophysiology and neuro-anatomy, drawing on evidence from across the cognitive sciences. There are many theoretical options available, and I have indicated some dimensions along which illusionist theories may differ. Some of the issues and positions will be similar to those discussed by conservative realists, but they will assume a new aspect once the commitment to realism is dropped, and we can expect new connections to appear and new theoretical options to present themselves.

Most people find it incredible, even ludicrous, to suppose that phenomenal consciousness is illusory. But if the illusion has been hardwired into our psychology for good evolutionary reasons, then that is to be expected. The question is not whether illusionism is intuitively plausible, but whether it is rationally compelling. If we had a detailed and well-supported illusionist theory, which fully explained our reports, judgments, and intuitions about our own consciousness, would we still want to insist, on reflection, that a hard problem remained? The best way to find out will be to try to construct such a theory.

Our introspective world certainly seems to be painted with rich and potent qualitative properties. But, to adapt James Randi, if Mother Nature is creating that impression by actually equipping our experiences with such properties, then she's doing it the hard way.<sup>16</sup>

---

<sup>16</sup> Earlier versions of this article were presented at The Open University and the University of Crete, and at a 'consciousness cruise' organized by Dmitry Volkoff and the Moscow Center for Consciousness Studies in June 2014, where Jesse Prinz presented a comment on it. My thanks to Jesse and to the audiences on those occasions, mentioning in particular Philip Goff, Martine Nida-Rümelin, Carolyn Price, and Michael Tye. Thanks are also due to Ned Block, Daniel Dennett, Eileen Frankish, Nicholas Humphrey, and Maria Kasmirli for their advice and suggestions. I am especially grateful to David

## References

- Blackmore, S. (2011) *Zen and the Art of Consciousness*, London: Oneworld Publications.
- Carruthers, P. (2000) *Phenomenal Consciousness: A Naturalistic Theory*, Cambridge: Cambridge University Press.
- Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford University Press.
- Chalmers, D.J. (2003) The content and epistemology of phenomenal belief, in Smith, Q. & Jolic, A. (eds.) *Consciousness: New Philosophical Perspectives*, pp. 220–272, Oxford: Oxford University Press.
- Dennett, D.C. (1988) Quining qualia, in Marcel, A.J. & Bisiach, E. (eds.) *Consciousness in Modern Science*, pp. 42–77, Oxford: Oxford University Press.
- Dennett, D.C. (1991) *Consciousness Explained*, New York: Little, Brown.
- Dennett, D.C. (2003) Who's on first? Heterophenomenology explained, *Journal of Consciousness Studies*, **10** (9–10), pp. 19–30.
- Dennett, D.C. (2005) *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*, Cambridge, MA: MIT Press.
- Dennett, D.C. (2007) Heterophenomenology reconsidered, *Phenomenology and the Cognitive Sciences*, **6** (1–2), pp. 247–270.
- Frankish, K. (2012) Quining diet qualia, *Consciousness and Cognition*, **21** (2), pp. 667–676.
- Hall, R.J. (2007) Phenomenal properties as dummy properties, *Philosophical Studies*, **135** (2), pp. 199–223.
- Harman, G. (1990) The intrinsic quality of experience, in Tomberlin, J. (ed.) *Philosophical Perspectives, Vol. 4, Action Theory and Philosophy of Mind*, pp. 31–52, Atascadero, CA: Ridgeview.
- Hill, C.S. (1997) Imaginability, conceivability, possibility and the mind–body problem, *Philosophical Studies*, **87** (1), pp. 61–85.
- Hill, C.S. & McLaughlin, B.P. (1999) There are fewer things in reality than are dreamt of in Chalmers's philosophy, *Philosophy and Phenomenological Research*, **59** (2), pp. 445–454.
- Howell, R. (2015) The Russellian monist's problems with mental causation, *The Philosophical Quarterly*, **65** (258), pp. 22–39.
- Humphrey, N. (1992) *A History of the Mind: Evolution and the Birth of Consciousness*, New York: Simon and Schuster.
- Humphrey, N. (2006) *Seeing Red: A Study in Consciousness*, Cambridge, MA: Harvard University Press.
- Humphrey, N. (2011) *Soul Dust: The Magic of Consciousness*, Princeton, NJ: Princeton University Press.
- Kripke, S.A. (1980) *Naming and Necessity*, new ed., Oxford: Blackwell.
- Levine, J. (2001) *Purple Haze: The Puzzle of Consciousness*, Oxford: Oxford University Press.
- Loar, B. (1990) Phenomenal states, in Tomberlin, J.E. (ed.) *Philosophical Perspectives, Vol. 4, Action Theory and Philosophy of Mind*, pp. 81–108, Atascadero, CA: Ridgeview.

---

Chalmers for his detailed comments on earlier drafts, from which the article has benefited considerably.

- Noë, A. (ed.) (2002) *Is the Visual World a Grand Illusion?*, Exeter: Imprint Academic.
- Papineau, D. (2002) *Thinking about Consciousness*, Oxford: Oxford University Press.
- Pereboom, D. (2011) *Consciousness and the Prospects of Physicalism*, New York: Oxford University Press.
- Place, U.T. (1956) Is consciousness a brain process?, *British Journal of Psychology*, **47**, pp. 44–50.
- Prinz, J.J. (2004) The fractionation of introspection, *Journal of Consciousness Studies*, **11** (7–8), pp. 40–57.
- Randi, J. (1997) Science and pseudoscience, in Terzian, Y. & Bilson E. (eds.) *Carl Sagan's Universe*, pp. 170–108, Cambridge: Cambridge University Press.
- Rey, G. (1992) Sensational sentences switched, *Philosophical Studies*, **68** (3), pp. 289–319.
- Rey, G. (1995) Towards a projectivist account of conscious experience, in Metzinger, T. (ed.) *Conscious Experience*, pp. 123–142, Exeter: Imprint Academic.
- Rey, G. (1998) A narrow representationalist account of qualitative experience, in Tomberlin, J.E. (ed.) *Philosophical Perspectives, Vol. 12, Language, Mind, and Ontology*, pp. 435–457, Boston, MA: Blackwell.
- Rey, G. (2005) Philosophical analysis as cognitive psychology: The case of empty concepts, in Cohen, H. & Lefebvre, C. (eds.) *Handbook of Categorization in Cognitive Science*, pp. 71–89, Amsterdam: Elsevier.
- Rey, G. (2007) Phenomenal content and the richness and determinacy of colour experience, *Journal of Consciousness Studies*, **14** (9), pp. 112–131.
- Searle, J.R. (1997) *The Mystery of Consciousness*, New York: The New York Review of Books.
- Smart, J.J.C. (1959) Sensations and brain processes, *The Philosophical Review*, **68** (2), pp. 141–156.
- Strawson, G. (2006) Realistic monism: Why physicalism entails panpsychism, in Freeman, A. (ed.) *Consciousness and its Place in Nature*, pp. 3–31, Exeter: Imprint Academic.
- Tartaglia, J. (2013) Conceptualizing physical consciousness, *Philosophical Psychology*, **26** (6), pp. 817–838.
- Tye, M. (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, Cambridge, MA: MIT Press.
- Tye, M. (2000) *Consciousness, Color, and Content*, Cambridge, MA: MIT Press.
- Wright, E. (2008) Introduction, in Wright, E. (ed.) *The Case for Qualia*, pp. 1–42, Cambridge, MA: MIT Press.