elizabethirv@gmail.com
mark.sprevak@ed.ac.uk

Eliminativism about consciousness

Liz Irvine
University of Cardiff

Mark Sprevak
University of Edinburgh

7 February 2018

1   Introduction

In this chapter, we examine a radical philosophical position about consciousness: eliminativism. Eliminativists claim that consciousness does not exist and/or that talk of consciousness should be eliminated from serious science. These are strong positions to take, and require serious defence. To evaluate these positions, the chapter is structured as follows. In Section 2 we introduce the difference between entity eliminativism and discourse eliminativism and outline the typical strategies used to support each. Section 3 provides a brief overview of the kinds of consciousness we refer to throughout the chapter. Section 4 focuses on entity eliminativist arguments about consciousness: Dennett's classic eliminativist argument (4.1), a rebooted version of Dennett's argument (4.2), and recent arguments for 'illusionism' (4.3). In Section 5, we examine discourse eliminativist arguments about consciousness: methodological arguments from scientific behaviourism (5.1), arguments based on the empirical accessibility of phenomenal consciousness (5.2), and a stronger position of discourse eliminativism aimed at both phenomenal and access consciousness (5.3). In Section 6, we offer a brief conclusion.

2   Eliminativism

If you meet an eliminativist the first question to ask is: Do you want to eliminate *entities* or *talk* about entities? For any given *X*, an eliminativist might assert either

or both of:

1. *X*s do not exist
2. We should stop engaging in *X*-talk, using *X*-concepts, or other practices putatively associated with *X* in serious science

We will call (1) *entity eliminativism* and (2) *discourse eliminativism*. Entity eliminativists claim that we should exclude a specific entity from the catalogue of entities assumed to exist. This may be a matter of removing a particular individual (e.g. Zeus) but it could also be a matter of removing a property (e.g. being phlogisticated), event (e.g. spontaneous generation), kind (e.g. ghosts), or process (e.g. extrasensory perception). In contrast, a discourse eliminativist seeks to eliminate certain ways of talking, thinking, and acting from serious science (e.g. talk about, and practices that attempt to investigate, gods, being phlogisticated, spontaneous generation, ghosts, or extrasensory perception).[1]

Entity and discourse eliminativism are distinct but obviously not unrelated. Existential status does not float free from what we say, think, and do in serious science. However, the relationship between the two is not so tight that entity eliminativism and discourse eliminativism should be collapsed into one claim.[2] Someone might endorse one form of eliminativism but not the other. An entity eliminativist claims that we should eliminate some specific entity (e.g. atoms) but may wish to preserve talk, thought, and practices associated with that entity in science. Mach claimed that atoms do not exist but that physicists should continue to engage in atomic talk, thought, and action for its predictive and heuristic benefits – the atom 'exists only in our understanding, and has for us only the value of a *memoria technica* or formula' (Mach, 1911, p. 49). A discourse eliminativist claims that we should eliminate specific

---

[1]In focusing on 'serious' science, the discourse eliminativist makes no claim about whether this talk or similar talk, thought, and practice should be eliminated from other aspects of human life. What might be unacceptable in serious science may be tolerated, or welcomed, in popular exposition of that science, folk tales, religious practice, jokes, or science fiction. The boundary between 'serious' science and other aspects of human life is not sharply defined, and for the purposes of this chapter, we do not attempt to define it. We simply identify 'serious' science as work currently recognised as such by the scientific community, in contrast to, say popular exposition of such scientific research, adaptation of that scientific research for other ends, or training that is merely propaedeutic to conducting scientific research.

[2]Quine (1980) offers a bridge from discourse eliminativism to entity eliminativism with his quantificational criterion of ontological commitment. However, this bridge does not link the two forms of eliminativism in a deductively certain way. It relies on numerous assumptions that are contentious in this context: about the aims of the scientific discourse, about the practice of stating truths having paramount value, and about the semantics of the discourse's terms. Notably, Quine only proposed his criterion for fundamental theories. Participants in this debate (about realism and eliminativism about consciousness) are unlikely to agree that the theories in question are fundamental.

ways of talking, thinking, and acting in science but might maintain that the entity at which they are directed nevertheless exists. In Section 5 we will see an example of this position in scientific behaviourism's treatment of conscious experience.

Entity eliminativism is defined by its departure from realism and agnosticism. A realist says, 'Xs exist', an agnostic says, 'We are not in a position to know whether Xs exist or not', and an eliminitivist says, 'Xs do not exist'. In order for an entity eliminativist to have a claim worthy of the name, she needs to have a genuine, not merely a verbal, disagreement with the realist and agnostic. To this end, the eliminativist needs to make some assumptions about Xs and those assumptions need to be shared with the realist and agnostic. The realist, agnostic, and eliminativist should agree on what Xs *would be like if they were to exist*. What they disagree about is whether this entity *does* exist. The realist says 'Yes' (or 'Probably yes'), the agnostic, 'Don't know', the entity eliminativist, 'No' (or 'Probably not'). Consequently, an argument for entity eliminativism involves two ingredients. The first is some way to identify the subject matter under dispute that is acceptable to realist, agnostic, and eliminativist. This is often done by providing a description of the essential properties of the entity. However as we will see in Section 4.2, this is not the only way to do it. The second ingredient is an argument to the effect that no such entity exists. If the entity was identified by description, the second step may be an argument to show that nothing satisfies this description. Mallon et al. (2009) summarise how this strategy has been applied to argue for entity eliminativism about beliefs: First, claim that in order for something to be a belief, it must satisfy a certain description D (given by folk psychology). Second, argue that nothing satisfies description D (because folk psychology is false). Third, since nothing satisfies D, conclude that beliefs do not exist.

In contrast to entity eliminativism, discourse eliminativism targets specific ways of talking, thinking, and acting in science. Let us say that the concept 'ding dongs' refers to nanoscale, spherical, tentacled lifeforms. A discourse eliminativist says that scientists should stop talking, thinking about, and pursuing research programmes about ding dongs. One motivation for this may be that there are no ding dongs (i.e. one is an entity eliminativist about them). But being an entity eliminativist is neither necessary nor sufficient for being a discourse eliminativist. One might think that ding dongs exist (or be agnostic about them) but nevertheless think that scientists should avoid talking and thinking about them because it is unproductive, misleading, or otherwise unhelpful. Conversely, one might be an entity eliminativist about ding dongs and reject discourse eliminativism about them. That is, one may believe that this talk, thought, and practice is useful to scientists even though no ding dongs exist: it may provide useful concepts (perhaps the ding-dong concept is a useful way to group lifeforms) or encourage useful practices (looking for entities at certain spatial scales).

Arguments for discourse eliminativism typically have a negative part and a positive part. The negative part aims to establish that the talk, concepts, and practices targeted for elimination are somehow unhelpful, damaging, misleading, or otherwise problematic for science. An eliminativist might argue that the relevant discourse is too subjective, hard to verify, does not generalise well, does not pick out a natural kind, produces intractable disagreements, does not cohere with other scientific talk, or otherwise leads to a degenerative scientific research programme. However, even if these negative points land, they are rarely sufficient in themselves to motivate discourse eliminativism. Science does not normally change course unless a better alternative is available. The positive part of the eliminativist's argument aims to show that an alternative, and better, way of conducting science avoids these faults. The principal claim of the discourse eliminativist is that a proposed alternative is, on balance, better for achieving the goals in science than that the ways of talking, thinking, and acting targeted for elimination. Diverse virtues may weigh in this argument, including purely epistemic virtues (e.g. telling the truth, not positing things that do not exist), but also including more predictive, pragmatic, theoretical, and cognitive virtues.

3 Consciousness

In this chapter we make use of Block's distinction between access consciousness and phenomenal consciousness (Block, 1990; Block, 2007). 'Access consciousness' refers to the aspects of consciousness associated with information processing, such as storage of information in working memory, planning, reporting, control of action, decision making, and so on. 'Phenomenal consciousness' refers to the subjective feelings and experiences that conscious agents enjoy: the feel of silk, the taste of raspberries, the sounds of birds singing. This is the 'feel-y', subjective, qualitative, what-it-is-like-ness, 'from the inside' aspect of consciousness. We also use 'qualia' to refer to this subjective feeling.[3] Following Frankish (2016a), we define 'experience' in a purely functional way to refer to mental states that are the direct output of sensory systems. Consequently, we do not assume that experience necessarily involves phenomenal consciousness.[4]

This chapter focuses on eliminativism about *phenomenal consciousness*. However, access consciousness will make an appearance in the final section (5.3). For the purposes of this chapter, we do not presuppose anything about the relationship

---

[3]This is how Dennett uses 'qualia' but departs from some authors who use 'qualia' to refer to non-representational aspects of conscious experience.

[4]Other authors (including Block) equate phenomenal consciousness with experience. We adopt Frankish's usage here for (slightly better) ease of exposition.

between access and phenomenal consciousness, though some of the eliminativist arguments discussed below do take a stand on this.

One of the most striking features of eliminativism about consciousness is that it is perceived as a philosophical position that is hard to motivate to the point of being self-evidently wrong. It is hard to doubt that we have phenomenal consciousness. Each of us appears to know by introspection on our own case that we have subjective feelings – and what is more, we appear to know this *for certain* in a way that does not leave room for rational doubt. Not even Descartes doubted the existence of his conscious experience. Yet the eliminativist does this. Critics of eliminativism have responded harshly. Frances writes, 'I assume that eliminativism about feelings really is crazy' (Frances, 2008, p. 241). Searle, 'Surely no sane person could deny the existence of feelings' (Searle, 1997). Galen Strawson agrees: eliminativists 'seem to be out of their minds', their position is 'crazy, in a distinctively philosophical way' (Strawson, 1994, p. 101). Chalmers says, 'This is the sort of thing that can only be done by a philosopher, or by someone else tying themselves in intellectual knots!' (Chalmers, 1996, p. 188). How can one apportion any credence to a view that denies an evident and unshakable truth about our mental life? (As we will see in Section 4, entity eliminativism is typically run in conjunction with an attack against the reliability of introspection.)

The discourse eliminativist about phenomenal consciousness faces a similar, although perhaps not quite so daunting, challenge. Discourse eliminativists seek to find problems with a discourse in science and offer a better alternative. The challenge a discourse eliminativist about phenomenal consciousness faces is that phenomenal consciousness appears to be incredibly important to human psychology. Humans care about the phenomenal feelings that accompany their experiences: about the feelings that accompany eating their favourite dish, scoring a winning in goal, being punched in the kidneys, or having their toes tickled. These feelings play a significant, although hard to pin down, role in their cognitive economy. For this reason, it seems that some reference to phenomenal consciousness *should* play a role in scientific psychology. A scientific psychology that never talked about phenomenal consciousness would be incomplete or inadequate in some way. Even if phenomenal feelings do not exist (as an entity eliminativist says), scientific psychology should still talk about phenomenal consciousness in order to explain why we (falsely, according the entity eliminativist) take such feelings to motivate us. Eliminating talk of phenomenal consciousness ignores an important aspect of the world and constitutes a failure of ambition for scientific psychology.

4 Entity eliminativism about consciousness

In this section, we examine three entity eliminativist arguments about phenomenal consciousness. The first is Dennett's 'Quining qualia' argument (1988). The second is our rebooted version of Dennett's argument that aims to avoid a standard objection (that Dennett mis-characterises phenomenal consciousness). The third is the recent research project of 'illusionism', which is related to Dennett's 'Quining qualia' argument but motivated on somewhat different grounds.

4.1 Dennett's eliminativism about qualia

Dennett's 'Quining qualia' argument appears to fit the template of a classic entity eliminativist argument: describe the essential properties of the supposed entity, show that nothing satisfies this description, and on this basis conclude that no such entity exists. The description that Dennett provides of phenomenal consciousness ('qualia') is that, in addition to having a phenomenal feel, qualia also are ineffable (not describable in words); intrinsic (non-relational); private (no inter-personal comparisons are possible); and directly accessible, e.g. via direct acquaintance. The final property is related to the idea that we have privileged, incorrigible, or infallible access to qualia. Dennett argues that nothing satisfies this description. As a result, 'Far better, tactically, to declare that there simply are no qualia at all' (Dennett, 1988, p. 44).

Dennett uses a number of 'intuition pumps' to get to this conclusion, which we summarise here.

First, it is plausible that how things phenomenally feel is bound up in how one evaluates or is able to categorise or discriminate between them. Someone's first taste of a particular wine may be different to how it tastes to them after having having become a wine aficionado. At first, my taste of the wine was bound up with judgements of yukkiness and an inability to easily tell one wine from another. My current taste of the wine is bound up in gustatory enjoyment and an ability to finely discriminate between different wines. This suggests that qualia are not intrinsic properties of experience: that the taste of this specific wine does have not have a particular conscious, qualitative feeling (quale) for me independently of how I evaluate or categorise it. Instead, the way that wine (and other things) consciously taste to me is at least partly determined by relational properties, such as whether I like it, and whether I can tell a Pinot Grigio from a Chardonnay.

This also puts pressure on qualia being directly accessible: it looks like I can't tell very much about my qualia from introspection. Say that as you get older, you start liking

strong red wine more. One possibility is that your sensory organs have changed, making strong reds taste different and more pleasurable compared to how they used to. On this scenario, you now have different qualitative feelings (more pleasurable ones) on tasting strong reds to those you had before (less pleasurable ones). Another possibility is that your sensory organs have stayed the same but your preferences for specific experiences has changed. You have roughly the same qualitative feelings but now you like those feelings more. On the first scenario your qualia change; on the second scenario your qualia stay the same. Dennett puts it to us that we would not be able to tell, merely from introspection, which scenario we are in. Yet this should be easy to do if we really had direct (or infallible or incorrigible) access to our qualia.

With respect to the putative ineffability and privacy of qualia, Dennett refers to Wittgensteinian arguments that render entirely private and incommunicable states as senseless. Dennett argues that there is nevertheless a sense in which experiences are practically ineffable and private, albeit just not in the 'special' way intended by qualia-philes. Imagine two AI systems that learn about their environment in a fairly unsupervised way and so develop different internal systems of categorising colour (adapted from Sloman and Chrisley, 2003). These two systems will end up with some states that are (at least to some degree) private and ineffable. One system's 'blue states' will be somewhat different to the other system's 'blue states' just in virtue of the internal differences in the systems (e.g. the 'blue states' of each system will be triggered by a slightly different range of hues). In the same way, humans can be in distinct (practically) ineffable and private states because of idiosyncrasies in their cognitive processing. Some differences one may discover empirically (for example, we discover that you and I disagree about whether a particular paint chip is blue), and so we can make our experience more 'effable', and less private. Dennett's point is that ineffability and privacy of experience only amounts to this: practical and graded difficulties in assessing which internal state we are in, not essential properties of our experience.

In light of Dennett's considerations, it looks like supposedly phenomenal experience does not satisfy the description of qualia: there is nothing ineffable, intrinsic, private and directly accessible that determines the way that things (phenomenally) seem to an experiencer. Whatever produces our judgements and reports about phenomenal experience, it is not an entity of the hypothesised kind. Qualia, as characterised above, do not exist.

A popular response to Dennett's argument is to deny that qualia are adequately characterised by his description. This would effectively undermine Dennett's argument at the first step. To this end, qualia realists nowadays tend to assume only that qualia have a *phenomenal character* (a 'what-is-it-like-ness' or subjective feeling)

(Carruthers, 2000; Levine, 2001; Kind, 2001; Tye, 2002). They do not require that qualia also be intrinsic, private, ineffable, or directly accessible. Frankish (2012) describes this as a switch in the debate from 'classic qualia' to 'diet qualia'.[5] While positing classic qualia may be controversial (perhaps for the reasons Dennett gives), positing diet qualia should not be:

> Philosophers often use the term 'qualia' to refer to the introspectively accessible properties of experiences that characterize what it is like to have them. In this standard, broad sense of the term, it is very difficult to deny that there are qualia. There is another, more restricted use of the term 'qualia', under which qualia are intrinsic, introspectively accessible, nonrepresentational qualities of experiences. In my view, there are no qualia, conceived of in this way. They are a philosophical myth. (Tye, 2002, p. 447)

Dennett does not wish to eliminate classic qualia in order to make room for diet qualia. He wishes to eliminate both classic qualia and diet qualia.[6] We have seen that Dennett's eliminativist argument is ineffective against diet qualia.[7] In the next section, we consider how Dennett's argument might be adapted to work against diet qualia.

4.2 Quining qualia rebooted

Dennett's 'Quining qualia' argument identifies qualia by description. Once we switch to diet qualia, descriptions appear to be little use as there is nothing to identify diet qualia apart from their (contested) phenomenal feel. So, rather than attempt to identify the target for elimination (or realism) by description, we need to identify it in some other way.

A common strategy is to identify qualia by a kind of ostension: consider specific

---

[5]A similar distinction is described by Frankish (2016a) between *weak* and *strong* illusionism, and by Levine (2001) between *modest* and *bold* qualophilia.

[6]"Philosophers have adopted various names for the things in the beholder (or properties of the beholder) that have been supposed to provide a safe home for the colors and the rest of the properties that have been banished from the "external" world by the triumphs of physics: "raw feels," "sensa," "phenomenal qualities," "intrinsic properties of conscious experiences," "the qualitative content of mental states," and, of course, "qualia," the term I will use. There are subtle differences in how these terms have been defined, but I'm going to ride roughshod over them. In the previous chapter I seemed to be denying that there are any such properties, and for once what seems so is so. I am denying that there are any such properties' (Dennett, 1991, p. 372).

[7]Frankish (2012) argues that diet qualia and classic qualia are not conceptually distinct and so Dennett's original argument works against both classic qualia and diet qualia. We will not consider his argument for this here.

examples of qualia and then generalise to the kind they have in common. One identifies the subject matter by asking one's interlocutor to consider those of her mental episodes that putatively have qualia (consider the feel of silk, …), draw her attention to the supposedly felt aspects of these episodes, and say *more aspects of mental life of this kind.* A set of examples, and how they are relevantly similar, thus are intended to fix the meaning of 'qualia'.[8] The realist and the eliminativist may agree on this strategy. They may agree on the set of examples and how we should take them to be relevantly similar (for example, they may have no difficulty in generalising to new cases). What the realist and eliminativist disagree about is whether the phenomenal feel that *appears* to be present in these cases *really is* present. The realist claims that in the examples those experiences have, or instantiate, a property, 'what-it-is-like-ness', which should be added to our ontology. This 'what-it-is-like-ness' or phenomenal character is a real property – as real as anything we know. The realist then says that explaining what this property is, how it comes about, and how it relates to physical properties is the job of a theory of consciousness. The entity eliminativist holds that no such property exists (or is instantiated in the relevant cases). According to her, the examples are, in a sense, deceptive: they appear to show instantiation of a property, but that appearance is mistaken. There is no such property of experience.

An entity eliminativist denies the existence of qualia, but she does not deny the existence of our judgements, beliefs, and desires about qualia. This allows her to agree with a lot of what a realist says about experience. She can agree that we are *disposed to believe* that our experience has qualia. She can agree that it is *hard for us to deny* that our experience has a 'what-it-is-like-ness' or phenomenal feel. She can agree that our judgements about qualia *motivate us to act* in congruent ways. Nevertheless, the eliminativist says that these judgements about qualia are false. Our beliefs about qualia are comparable those of the ancient Greeks about Zeus: deeply held and capable of motivating action, but false. We should no more take on the tasks of explaining what qualia are, how they arise, and how they relate to physical properties than we should for Zeus.

Our rebooted version of Dennett's argument is divided into two steps. The first

---

[8]A range of specific examples appear to play this role in Chalmers (1996), Chapter 1. Schwitzgebel (2016) outlines a similar strategy, although with the commitment that there should be a 'single obvious folk-psychological concept or category that matches the positive and negative examples' (We do not think that either the realist or eliminativist need admit this). Nida-Rümelin (2016), Section 3 outlines a similar strategy to identify 'experiential' properties, although she argues that if this strategy works, there can be no possibility of failure to refer so eliminativism is precluded. The general strategy of reference fixing by ostending examples from a single kind follows that of a causal theory of the reference, although in this case the subject's relation to the examples need not be causal (e.g. it could be some sort of non-causal acquaintance relation).

step aims to establish an epistemic claim: we do not know which (diet) quale our experience instantiates. This claim is meant to be stronger than a failure of infallibility or incorrigiblity: it is that we lack *any* knowledge at all of which quale our experience currently instantiates. The second step leverages this scepticism to argue for qualia eliminativism. If the instantiation of one quale rather than another is unknowable (as the first step says), then such facts are differences that make no difference to the rest of the world, and on that basis, they should be eliminated.

As a starting point, consider that sometimes it is hard to tell which subjective feeling you have. Slow or subtle shifts in experience may leave you uncertain about the identity of your feelings – is the feeling you have now on looking at a blue painting the same as that you had one minute ago? However, such cases may only provide us with epistemically 'bad' cases of qualia knowledge, where for some reason we are not in a strong position to know which quale we currently have. Showing that there are 'bad' cases does not show that we can never know which quale we have, and cannot sustain an eliminativist argument of the form just described.[9]

Focus instead on 'good' cases of qualia knowledge, where we appear to know, with absolute certainty, whether our qualia are the same or different between two conditions. Such cases often involve sudden or dramatic changes in experience between the two conditions. For example, if a blue painting in front of you were instantaneously changed to a yellow painting, you would know, not only that the painting had changed but also that your accompanying *phenomenal experience* had changed (perhaps you regard the latter as part of your evidence base for the former). Colour inversion thought experiments provide particularly stark cases of such qualia knowledge. The rebooted version of Dennett's argument aims to show that even in such maximally 'good' cases, one has no knowledge about which quale one has.

Suppose that while Lara is asleep a neurosurgeon operates on her brain. On waking, Lara finds that the world looks different to her: objects that looked blue now look yellow. No one else notices a change so Lara concludes that something must have happened to her. *Prima facie*, Lara appears to know that her qualia have changed. She knows because 'things look different'. However, Lara's subjective evidence is compatible with two hypotheses:

(Q) Lara's qualia have changed from those she had yesterday
(R) Lara's qualia are the same but her memories of past colour experiences have changed

The problem is that any world in which Q is true would be introspectively indistin-

---

[9]It is unfortunate that many of Dennett's intuition pumps involve subtle and slow changes in qualia. This has focused attention on failures of infallibility and incorrigibility in 'bad cases'. The more worrisome lesson is that no knowledge of qualia is possible even in epistemically 'good' cases.

guishable to Lara to a world in which R is true. On the basis of introspection she can tell that 'things look different', but she cannot tell what is responsible: a change in her memory, a change in her current qualia, or a mix of both. The problem is more serious than a failure of certainty or infallibility. Q and R are equally supported by her introspective evidence. Introspection appears to give Lara *no* knowledge at all about whether her qualia have changed.

Lara has access to other sources of knowledge than introspection. What if she looks at changes in her brain across the two conditions? For the sake of argument, put Lara into the strongest possible epistemic position with regard to any physical changes in her brain. Lara has knowledge of a completed neuroscience and full scans of her brain before and after surgery. Furthermore, suppose (unrealistically) that there is a clear separation in the neural basis of Lara's sensory systems and her memories of past experiences and that this is known to her. She can then reason as follows: 'If the scans reveal my surgery affected only my sensory channels and left my memory system intact, I have reason to favour Q over R because only my current sensation and not my memories of past experiences have been affected. Conversely, if the scans reveal that my surgery affected only my memory systems and left my sensory channels intact, then I have reason to favour R over Q because only my memories and not mechanisms that supports my current sensations have been affected.' Thus it seems that empirical evidence can do for Lara what introspection cannot: give her evidence about the identity of her current quale. (Of course, it is possible that both Lara's sensory and her memory systems are affected by the surgery, but ignore this possibility as it would not help her).

Unfortunately, Lara's epistemic situation is worse than it might appear. Brain scans and neuroscientific knowledge provide her with information about changes to *sequences of neural events* but this only helps Lara with Q and R if she knows *where* in that causal chain her qualia experience occurs. Lara does not know this and it is not part of the concept of diet qualia. Consider two possibilities:

1. Sensation → *qualia experience* → memory access
2. Sensation → memory access → *qualia experience*

Above we assumed something like (1): Lara's qualia experience occurs after her current sensation and before memory access. If (1) is true, then Lara's reasoning holds. Her memory access is causally downstream from her qualia experience and so changes to her memory systems should not affect her current qualia experience. Conversely, changes to her sensory system (e.g. swapping her sensory channels) would affect her current sensation but not her ability to access memories of past experiences. We have causal inference here – which falls short of deductive certainty (as a cause could conceivably produce any effect) – but at least Lara has some reason to prefer Q or R.

However, if the causal order is described by (2), then Lara's inference fails. A surgical change to her memory system could equally well result in a change in her current qualia or a change in her memory of past experiences, or both. A surgical change to her sensory system could equally well produce a change in her current qualia or a change to the outputs of her memory systems, or both. The two factors (memory and current experience) are confounded.

The problem is that no one knows whether (1), (2) or any number of other proposals about the location of qualia in the causal order is correct. This is not a limit of technology or knowledge of neuroscience – a better scanner or more neural information would not help. Nor is it something with which the concept of diet qualia can help. Conceivably, Lara's introspective reports could be correlated with her neural events to find out where in that causal order her experience of qualia falls. However, we have seen that there is no reason to trust her introspective reports in this case.[10] Neither introspection nor empirical knowledge nor some combination of the two provides Lara with knowledge about which quale she has. Even in this apparently 'good' case with an apparently clear contrast between the two conditions, Lara has no reason to favour Q or R, or to think that her current experience instantiates any other quale ('blue', 'yellow', 'tickle', 'hot'). Given our own current epistemic position, for all we know our own qualia may change all the time without us being able to detect it.[11]

This accounts for Step 1 of the argument. Step 2 claims that if the identity of one's current quale is not knowable (either via introspection or methods available to empirical science), then we should eliminate qualia from our ontology. From Step 1, we know that the specific quale instantiated in current experience cannot be detected. This suggests that qualia are a kind of extra 'wheel' that does not turn anything detectable in our ontology. Qualia have no characteristic, detectable effects. For if they did, we could use that to detect them. Their effect on us is always mediated by, and confounded with, other factors (such as memory). Therefore, positing qualia as independent, *sui generis* entities in our ontology seems unmotivated. Like Wittgenstein's 'beetle',

> [This] thing in the box has no place in the language-game at all; not even as a something; for the box might even be empty – No, one can 'divide through' by the thing in the box; it cancels out, whatever it is. (Wittgenstein, 1958, §293)

It is of course open to a realist to insist that qualia should be included in our ontology

---

[10]Other problems with such studies are described in Section 5.2.

[11]Dennett (2005), Chapter 4 presents a similar argument for eliminating qualia using the phenomenon of change blindness that again relies on cross-time comparisons (for all we know, our colour qualia may be changing all the time without us noticing).

irrespective of our ability to detect them (similarly, a beetle realist could stick to their guns). But if this is what realism amounts to it seems little more than dogma. In light of the confounds above, qualia do not appear to earn their keep. Better to say, that as needless ontological cruft, they should be eliminated.

Unlike Dennett's argument, this argument does not rely on the assumption that qualia are intrinsic, private, ineffable, or directly accessible. Before closing, we wish to flag two problems with the argument.

First, one might wonder, even if the argument is correct, why it still *seems* to us that qualia exist. This seeming does not go away even if one accepts the eliminativist conclusion. On this basis, one might press for a residual role for qualia that provides more than an eliminativist allows: more than merely having a set of dispositions to make judgements about which qualia we have, or having beliefs about qualia (both compatible with such judgements and beliefs being mistaken). Our relationship to qualitative aspects of experience is arguably more primitive than this. It *seems* to us that our experiences have qualia and this appearance is the evidence for our beliefs about qualia. But how can this impression, this pre-doxastic 'seeming', be produced? One might tell a mechanistic and adaptive story about how humans arrive at deeply held false beliefs (Dennett, 1991; N. Humphrey, 1992). But what mechanistic story can be told that explains the production of *seemings* that generate and appear to confirm these beliefs? This is the 'illusion problem' discussed in the next section and it remains an unsolved challenge for qualia eliminativists.

Second, one might question whether the qualia scepticism described by Step 1 is sufficiently motivated. Step 1 relies on a memory-based comparison: Does Lara know whether her quale today is the same as that yesterday? A qualia realist might concede that she does not know this (perhaps because of confounds with memory), but deny that Lara lacks any knowledge at all about which quale her experience instantiates. Imagine looking on a mountain scene with green grass, grey rock, and blue sky. Multiple qualia seem to be simultaneously instantiated in your experience (*what-it-is-like to see green*, *what-it-is-like to see grey*, *what-it-is-like to see blue*, and so on). You appear to be able to tell the difference between these qualia (you can make similarity judgements, detect that there are lots of such feelings versus few, and so on). None of these judgements appear to rely on memory. Within the domain of current experience therefore, you appear to have *some* knowledge about qualia identity and difference. But then, why think that current qualia are always confounded in their effects with memory and so are idle wheels that turn nothing?

4.3 The illusionist movement

Recently, interest in eliminativist approaches to phenomenal aspects of consciousness has been rekindled by Keith Frankish, in particular in a special issue of *Journal of Consciousness Studies*. Frankish outlines 'illusionism' as the view that experiences have no phenomenal properties and so that phenomenal consciousness is 'illusory'. We think we have experiences with phenomenal properties, but in fact we do not. Illusionism is a form of entity eliminativism about phenomenal consciousness even if the label 'eliminativism' is avoided for rhetorical reasons. It is motivated somewhat differently to Dennett's entity eliminativism, and has a slightly different focus, so is worth discussion in its own right.[12]

First, illusionism is partly motivated by taking seriously the idea that phenomenal properties, and phenomenal consciousness, cannot be accounted for scientifically. Illusionism is seen as a way out of this problem. Second (and relatedly), the reasons for favouring illusionism are mainly rather general, theoretical reasons. The theoretical virtue of simplicity, or conservativism, suggests that the fewer entities/properties the better. Since illusionism gets rid of the metaphysically and epistemically problematic phenomenal properties, illusionism is better than alternative realist positions. Third, illusionism is often argued to be a research program rather than a set of worked-out claims, and this research programme is worth pursuing more than its alternatives. As we will see, illusionism comes with a range of difficult open questions.

Illusionism follows a slightly different tack to the typical argument for entity eliminativism described in Section 2. The first step is supposed to be identifying the contested entity/property in a way that can be generally accepted. This is not straightforward for phenomenal properties (see discussion both above and below). Second, the arguments motivating illusionism are not direct arguments to the effect that phenomenal properties, as described, do not exist; the position is largely motivated on other grounds (e.g. theoretical simplicity). The third step is to conclude that phenomenal properties do not exist: this is concluded by proponents of illusionism, but one could arguably also treat illusionism as a promising research program without in advance committing to this conclusion.

Challenges to illusionism come in roughly three forms (the first two roughly track two of the steps above).

---

[12]Dennett-style eliminativism treats our ontological commitment to phenomenal consciousness as a theoretical mistake: there is nothing that satisfies the description of qualia, or qualia are ontologically inert and therefore it is safe to eliminate them. Somewhat differently, one can see illusionism as treating our ontological commitment to phenomenal consciousness as an introspective or perceptual mistake: we 'perceive' (via introspection) that we experience phenomenal properties but we do not (hence the illusionism title). However, see Frankish (2016b) for ways of blurring the boundary between theoretical and introspective mistakes.

First, one can argue that it is not obvious and universally accepted what phenomenal consciousness is, or what phenomenal properties are, such that a proposal to eliminate them is comprehensible. Mandik (2016) states that 'phenomenal' is a technical (not folk) term, but one that is not clearly defined. As such, both eliminativist and realist talk about 'phenomenality' is unwarranted; in neither case is there a clear target to be eliminativist or realist about. Schwitzgebel (2016) tries to provide a minimal 'definition by example' that is not committed to any particular (troublesome) metaphysical or epistemic commitments, but as Frankish (2016b) points out, this is not substantive enough to sway the debate one way or the other.

Second, one can reject some of the main theoretical motivations for favouring illusionism as the best or most reasonable philosophical position available. For example, Balog (2016) defends the phenomenal concept strategy, which accepts the existence of the explanatory gap but preserves realism about phenomenal properties. Prinz (2016) defends a realist account of phenomenal properties, but one that tries to close the explanatory gap by providing neuroscientific explanation of at least some aspects of phenomenal consciousness. More generally, unless one is convinced that the hard problem or the explanatory gap present insuperable difficulties, one is unlikely to persuaded of illusionism.

Third, there are worries about how to solve the illusion problem. The illusion problem is how to account for the illusion of phenomenality: how one could think that one was having experiences that appear to have phenomenal properties without any phenomenal properties existing. Frankish (2016a) labels 'quasi-phenomenal properties' those physical properties (perhaps highly disjunctive and gerrymandered) that our introspection misrepresents as having phenomenal qualities. Quasi-phenomenal redness is, for example, the physical property that typically causes (false) representations of phenomenal redness in our introspective experience. According to Frankish, it is the tokening of these false representations that is responsible for the illusion of phenomenal consciousness. He likens the effect of these representations to that of other persistent, false representations in our perceptual system such as our representation of impossible figures like the Penrose triangle (N. Humphrey, 2011) or our representations of colours as 'out there' in the world (Pereboom, 2011).

The worry is how exactly this is supposed to work. It is not clear how false representations caused by non-phenomenal properties could produce an appearance or 'seeming' of phenomenality. And as Prinz puts it, '… what is it about beliefs in experience that causes an illusion of experience?' (2016, p. 194): how and why is it that these representations in particular cause such illusions when other sorts of representations do not? Related to this worry is a worry about how the false representations get their content (Balog, 2016). Representations of phenomenal feelings

are not like other empty representations ('Pegasus', 'largest prime') which appear to get their content by being built out of representations that do refer ('winged', 'horse', 'largest', 'prime'). How do false representations of phenomenal experience get their content if there are no phenomenal feels?

Despite illusionism promising to get away from the hard problem, Prinz (2016) argues that the illusion problem and the hard problem in fact face similar difficulties. In both cases we need to identify what phenomenal properties and phenomenal consciousness are. In the hard problem, we need to explain how they come out of 'mere matter': how phenomenality arises from an apparently non-phenomenal system. In the illusion problem, we need to explain how (vivid!) illusions of phenomenality come about in entirely non-phenomenal systems. The problem is to explain how an illusion of phenomenality (worthy of that name) arises in a non-phenomenal system. In both cases then, one needs to explain how something suitably like phenomenality arises from 'mere matter'. Yet by the time one has done this, it might be just as easy to be a realist as an illusionist.

Frankish (2016a) considers the relationship of illusionism to discourse eliminativism: 'Do illusionists then recommend eliminating talk of phenomenal properties and phenomenal consciousness? Not necessarily.' (p. 21). However, he suggests that discourse eliminativism could only be avoided by an illusionist if phenomenal terms in science were redefined to refer to quasi-phenomenal properties – the physical properties that typically cause the relevant false representations. This seems to us neither necessary nor likely.

First, as Frankish says, it would be a substantial departure from what these terms mean in others contexts, and so invite confusion. Second, although we agree that scientific psychology would need a way to keep track of the quasi-phenomenal properties, this would most naturally be done with a response-dependent characterisation: group together the physical properties that give rise to a specific (false) phenomenal representation and type a quasi-phenomenal property accordingly. Keeping track of quasi-phenomenal properties does not require redefining what phenomenal terms mean in science. Third, as Frankish (2016b) says, it is no part of illusionism to say that consciousness is not an important or useful illusion. Graziano (2016) and Dennett (1991) argue that phenomenal consciousness plays a central, adaptive role in our mental lives. It is reasonable to expect a scientific psychology would want to study this and that could be done while bracketing questions about the existence of phenomenal properties.[13] Similarly, a psychology that studied childhood dreams might want to talk about the role of representations of Santa Claus and unicorns – without attempting to re-define those terms to refer to physical entities,

---

[13]Dennett (1991)'s heterophenomenology provides one model of how an illusionist might avoid discourse eliminativism.

or positing real entities corresponding to those terms. Talk of phenomenal feels will likely remain in scientific psychology, albeit with the qualification that the entities that appear to stand behind this talk do not exist.

5 Discourse eliminativism about consciousness

We now turn to discourse eliminativism. Discourse eliminativism seeks to remove talk, concepts, and practices regarding phenomenal consciousness from science (even if phenomenal consciousness is still admitted to exist). In this section, we look at three discourse eliminativist arguments. The first is based on considerations raised by psychologists at the start of the twentieth century. The second is based on more contemporary concerns about how to study phenomenal consciousness independently of access consciousness and the mechanisms of reportability. The third is based on the worry that the concept of consciousness fails to pick out a scientifically usable category of phenomena.

5.1 Scientific behaviourism

One of the goals of scientific psychology in the first half of the twentieth century was to redefine psychology, not as the study of the mind, but of observable behaviour. Scientific behaviourists argued that scientific psychology should avoid talk of internal mental states, and in particular talk of conscious states (Hull, 1943; Skinner, 1953; Watson, 1913).

The rise of scientific behaviourism was at least partly due to the perceived failure of an earlier attempt to scientize psychology via trained use of introspection (Titchener, 1899). The debate on the nature of imageless thought was held up as an example of how unproductive that research programme was. One side in the debate appealed to introspection to argue that all thoughts were analysable into images; the other side used similar evidence to argue that some thoughts are imageless. The disagreement was widely seen as unresolvable because the evidence from disagreeing enquirers could not be compared in an unbiased way. By the mid-twentieth century, introspective methods were discredited and scientific study of conscious experience largely abandoned (G. Humphrey, 1951).

Scientific behaviourists sought to reposition psychology to avoid these methodological difficulties. The subject matter of scientific psychology should be publicly observable, publicly verifiable, or independently experimentally controllable events. Psychology should eliminate talk of conscious experience and use of introspective methods. However, this did not mean that behaviourists thought that internal

mental states, including states of consciousness, did not exist: 'The objection to inner states is not that they do not exist, but that they are not relevant in a functional analysis' (Skinner, 1953, p. 35).[14] Scientific behaviorists proposed an alternative way of talking, thinking and acting that they argued was superior (in predictive, explanatory, and methodological terms) to a scientific psychology that appealed to, or attempted to study, conscious experience. Phenomenal consciousness, irrespective of its ontological reality, should be excluded from scientific psychology because its study was methodologically flawed and appeal to it was unnecessary to explain behaviour.

Based on related considerations about verification and public accessibility, positivistically-inclined philosophers argued that various ontological and/or semantic lessons follow about conscious experience (Ryle, 1949; and less clearly, Wittgenstein, 1958). This prompted them to redefine mental state language in terms of behavioural dispositions and/or to eliminate qualitative conscious feelings from ontology. However, to connect these two lines of thought – one about scientific practice and the other about ontology/semantics of ordinary language – requires accepting auxiliary claims about verifiability, the role of science, and the scope of human knowledge. Such links are widely questioned today. Many scientific behaviourists did not perceive such links either and they argued for the elimination of talk of conscious experience from science based on pragmatic rather than metaontological concerns.

5.2   Eliminativism via independent access

A different methodologically motivated form of discourse eliminativism about phenomenal consciousness is found among some consciousness researchers today. This stems from the problems involved in trying to operationalise consciousness, or finding ways to experimentally probe it. One typical way of operationalising consciousness is via some kind of reportability: a subject is conscious of a stimulus if and only if they report it or respond to it in some way. This is fairly straightforward, but there are potential problems in using reportability as a marker of the presence of phenomenal consciousness, rather than as a marker of the cognitive capacities associated with consciousness. This can motivate a position of discourse eliminativism about phenomenal consciousness.

First, consider the distinction between phenomenal and access consciousness. Phenomenal consciousness refers to felt conscious experiences, (diet) qualia, raw feels, and so on. Access consciousness refers to the aspects of consciousness that are

---

[14]See Hatfield (2003) for discussion of the views of other behaviourists.

associated with or can be used in cognitive capacities like reasoning, action, verbal report, and so on. If we somehow knew that access and phenomenal consciousness were always bound together (no cognitive access without phenomenal consciousness and vice versa), then scientific ways of probing access consciousness would also function as scientific ways of probing phenomenal consciousness. That is, if phenomenal consciousness and aspects of consciousness related to cognitive access always go together, then probing these cognitive capacities just is to probe phenomenal consciousness. In this case (absent any other problems), it would be perfectly legitimate for the term 'phenomenal consciousness' to figure in scientific discourse, because the phenomenon it picks out is scientifically accessible.

The problem is that it is not obvious whether the aspects of consciousness picked out by access and phenomenal consciousness are always co-present. According to Block (1995), there may be instantiations of phenomenal consciousness (raw feels) without any related cognitive access (ability to respond or report about these raw feels). Block has outlined a number of examples where this might happen, including when subjects may have highly detailed and specific phenomenal experiences, but be unable to report the details of them (Sperling paradigm), cases of phenomenal consciousness of unattended items, and possibly cases of hemi-spatial neglect, where subjects do not appear to have phenomenal experiences from some part of their visual field (see Block, 2007; Block, 2011; Block, 2014; Irvine, 2011; Phillips, 2011, for discussion of some of these cases). In most of these cases, there is evidence that subjects are at least processing sensory information that they are unable to report about. Block's claim is that there is also a layer of untapped and unaccessed phenomenal consciousness present in these cases, in addition to whatever can be overtly reported or measured.

The lack of a way to probe the phenomenal aspect of consciousness independently of the accessibility aspect makes it difficult (or impossible) to scientifically assess these claims. It looks like any way of probing phenomenal consciousness requires that the experience had some measurable effect on the subject, possibly such that she can report it in some way. That is, accessing phenomenal consciousness relies on it being associated with some kind of cognitive function or capacity, so accessing phenomenal consciousness relies on it being associated with access consciousness. So, if an instance of phenomenal consciousness is not associated with access consciousness, then it looks like it we cannot tell if it is present or not. As Dehaene et al. (2006) note, whether participants in an experimental situation '… actually had a conscious phenomenal experience but no possibility of reporting it, does not seem to be, at this stage, a scientifically addressable question' (p. 209).

Partly in response to this worry, Block, Lamme and colleagues have argued for the possibility of *indirectly* investigating these purported instances of phenomenal

consciousness without accessibility (Block, 2011; Block, 2014; Lamme, 2006; Sligte et al., 2010). The idea here is to find some reasonable and measurable marker of the presence of consciousness in cases where phenomenal (and access) consciousness is clearly present (call this marker, M). This could be a particular neurophysiological signature (e.g. evidence of strong feed-forward processing), or a particular behavioural marker (e.g. ability to complete a particular type of task based on a set of visual stimuli). One then argues that if the special marker M is present in a subject, then regardless of whether the subject appears to be conscious of the test stimulus according to other standard measures of (access) consciousness, the subject is phenomenally conscious of the test stimulus. That is, special marker M guarantees that a subject is phenomenally conscious of the test stimulus, even if they don't report seeing it, or can't perform a range of actions that we usually associate with being conscious of a stimulus. The subject experiences phenomenal consciousness of the stimulus without having cognitive access to that experience.

However, problems of interpretation abound here. Such behavioural and neurophysiological evidence could be taken as indirect evidence of phenomenal consciousness without access consciousness, but it could also be interpreted as evidence of unconscious processing (if we got the special marker M wrong), or of graded cognitive access and phenomenal consciousness of the stimulus (see replies to Block, 2007). There are no direct scientific grounds on which to choose between these interpretations, because there is no direct way to assess whether marker M has anything to do with phenomenal consciousness.

One response to these discussions is to advocate discourse eliminativism about phenomenal consciousness. This is based on accepting that there is no direct way to probe phenomenal consciousness independently of cognitive access, and that there are no straightforward empirical ways of testing the claim that phenomenal consciousness can be present independently of cognitive access. In this case, the only aspect of consciousness that can definitely be probed scientifically is cognitive access, i.e. access consciousness. In terms of achievable scientific practice, the safest methodological route is to drop talk of phenomenal consciousness. Something like this position appears to be taken by a number of consciousness researchers (possibly including Dehaene).

This position is compatible with a range of positions about the ontology of phenomenal consciousness. One might be either an entity eliminativist or entity realist about phenomenal consciousness. One might claim that phenomenal consciousness can (possibly or probably) exist without cognitive access, or be agnostic about this possibility. Alternatively, one might argue, with Cohen and Dennett (2011) that if a phenomenally conscious state is not accessible to scientists or the subject having it (e.g. via some kind of report), then it is (evolutionarily, cognitively) implausible

to call it a state of consciousness at all. In this case, if phenomenal consciousness exists, it always co-occurs with cognitive access.

5.3 Eliminativism via identity crisis

The kind of discourse eliminativism about phenomenal consciousness outlined above is based on a problem about *accessing* the phenomenon in question. Another kind of discourse eliminativism about consciousness is based on the problem of *identifying* the phenomenon in question. For now, assume that phenomenal consciousness always co-occurs with access consciousness (perhaps for the reasons suggested by Cohen and Dennett above), so that we can (for the minute) just work with the term 'consciousness' which will pick out both. However, even with this problem out of the way, it is still questionable whether the concept of consciousness picks out a clear category of phenomena that is scientifically useful. If it does not, this provides a new motivation for discourse eliminativism about consciousness, and (by assumption) discourse eliminativism about phenomenal consciousness too.

Above, it was suggested that there is a reasonably broad consensus that assessing the presence or absence of consciousness has something to do with reportability. However, reportability can be realised in a number of ways some of which are incompatible with each other (see Irvine, 2013, for review). One 'objective measure' of consciousness from psychophysics relies on forced-choice tasks, where subjects are typically shown a masked stimulus for a very short period of time and are 'forced' to choose between two response options (e.g. stimulus present/absent, stimulus was a square/circle). On the basis of these responses, the subject's underlying 'sensitivity' to the stimuli is calculated. The resulting objective measure is highly stable and not subject to problematic biases, but it is liberal, and often attributes visual consciousness of stimuli to subjects who explicitly deny having any. As a result, it is sometimes criticised as being merely a measure of sensory information processing (e.g. Lau, 2008). Despite being acknowledged as problematic, objective measures such as these tend to be used in studies of visual consciousness because of their desirable properties as scientific measures (they are stable, bias-free).

In contrast, 'subjective measures' of consciousness tend to use free reports or a wide range of possible responses, sometimes generated in advance by the subjects themselves. The experimental methodology may be based around emphasising to subjects careful use of introspection, assessing subject's confidence in their reports (sometimes using wagering), or just recording simple, untutored responses. Subjective measures get much closer to what subjects themselves acknowledge about their conscious experience. However, the precise ways that subjective measures are generated can have a significant impact on whether consciousness is deemed to be

present or absent (or somewhere in between) (Sandberg et al., 2010; Timmermans and Cleeremans, 2015). As scientific measures, they are highly unstable and subject to bias. They also regularly conflict with objective measures of consciousness (except under artificial training conditions), and are generally accepted as being conservative (they do not normally capture all instances of conscious experience).

The difficulties in identifying an adequate measure of consciousness re-appear in debates about the neural correlate(s) or mechanism(s) of consciousness. The difficulty is that behavioural measures are key in identifying these correlates and mechanisms. Roughly speaking, one chooses a measure, identifies the neural activity that occurs when the measure says that consciousness is present, and treats this as 'the' mechanism of consciousness. However, using different behavioural measures (unsurprisingly) leads to the identification of different neural correlates. These span from 'early' neural activity for some liberal behavioural measures (which may capture early sensory processing), to 'late' and attention-based neural activity for conservative behavioural measures (which may capture later cognitive uptake of the conscious experience) (see Irvine, 2013). With no settled agreement on what counts as an adequate measure of consciousness, there can be no agreement on what the neural correlates and mechanisms of consciousness are.

This plethora of measures and mechanisms is not necessarily problematic in itself, but Irvine (2012) argues that there is no methodologically viable way of resolving problems about how to adjudicate between them when they conflict. Each measure has its pros and cons, but none exists that is both scientifically adequate (i.e. fairly stable over repeated measures and bias-free), and that fits with pre-theoretic commitments about what consciousness is. To choose one measure would be to (operationally) define consciousness by fiat, which would undermine the motivations for engaging in 'real' consciousness science in the first place. Furthermore, the mechanisms that correlate with these varied measures to do not form a well demarcated scientific kind, or even a well demarcated group of kinds: they have no more in common than any arbitrary group of mechanisms of perception and cognition. They range over mechanisms of sensory processing, attentional processes, and processes related to decision making, report, and meta-cognition.

This provides one set of reasons for eliminating talk of consciousness in serious science. There are a wide range of incompatible things that 'consciousness' could pick out, and no methodologically acceptable way of deciding between them. If a scientific concept is surrounded by serious methodological problems, then (if they are bad enough), this is motivation for eliminating the concept from science. This consideration is reinforced by pragmatic concerns in favour of eliminating talk of consciousness. Given that it is unclear what 'consciousness' refers to, this generates unproductive debates and miscommunication, it blocks the generation of useful

predictions and generalisations, and it promotes misapplications of research methodologies and heuristics. That is, there are negative consequences from continued use of the concept 'consciousness' in scientific practice.

There is also a better alternative. This alternative demands that researchers use terms that clearly demarcate the phenomena under study, potentially by referring to how they are experimentally operationalised. This could done by splitting up the phenomena by how they are measured (e.g. forced choice tasks, confidence ratings, or free report). Using these more specific terms avoids the problems above. By precisely specifying what the phenomenon is and how it is measured, there is no ambiguity about what phenomenon is picked out. This also makes it possible to identify the mechanism that generates it, make robust predictions and generalisations about the target phenomenon, and avoid miscommunication.

As before, discourse eliminativism is not tied to entity eliminativism (Irvine (2012)'s position does not entail entity eliminativism of any sort). Discourse eliminativism is solely about scientific representations, concepts, methods and practices, and what it is scientifically appropriate and useful to take as a target of study. Whatever 'consciousness' (access or phenomenal) refers to may be out there, even if it is not a useful scientific concept.

6   Conclusion

In this chapter we have reviewed a range of arguments for positions of entity and discourse eliminativism, primarily aimed at phenomenal consciousness. Entity eliminativists deny the existence of phenomenal consciousness, while discourse eliminativists deny the scientific utility of talking about phenomenal (and perhaps access) consciousness.

Entity eliminativism can be pursued in a number of ways. One standard method is to describe the entity in question, and then show that nothing exists that satisfies that definition (4.1). This can be expanded to the method of using examples to fix the subject matter, and generalising to the relevant entity or property (4.2). A third approach taken by illusionists (4.3) is use a loose definition of the relevant entity/property, but to argue that whatever this refers to, it is theoretically and metaphysically simpler and more productive to work on the assumption that it does not exist. One problem that faces entity eliminativists of all types is to solve the 'illusion problem', which is the parallel of the hard problem faced by realists about phenomenal consciousness. This is the problem of how something non-phenomenal can give rise to something that seems phenomenal, or fools us into thinking that there are phenomenal properties.

Positions of discourse eliminativism are also motivated in slightly different ways, though all focus on scientific methodology. Classic behaviourism focused on what could be unequivocally measured in a public and 'observable' way, eradicating talk of mental states (5.1). More recent scientific work on consciousness has tended to move away from discussion of phenomenal consciousness on the basis that it is not clear if scientific methodology can probe it independently of the cognitive abilities associated with access consciousness (5.2). An argument can be made that the general concept of consciousness should be eliminated from scientific talk given the problems in clearly demarcating the phenomenon in question (5.3). Eliminating discourse about phenomenal consciousness from science sounds like it takes away a key concept in explaining human behaviour. However, this is not necessarily the case: specific reports and judgements about phenomenal consciousness can still function in explanations, and as explanatory targets in their own right.

Acknowledgements

Bibliography

Balog, K. (2016). "Illusionism's discontent". In: *Journal of Consciousness Studies* 23, pp. 40–51.

Block, N. (1990). "Consciousness and accessibility". In: *Behavioral and Brain Sciences* 13, pp. 596–598.

— (1995). "On a confusion about a function of consciousness". In: *Behavioral and Brain Sciences* 18, pp. 227–247.

— (2007). "Consciousness, accessibility, and the mesh between psychology and neuroscience". In: *Behavioral and Brain Sciences* 30, pp. 481–548.

— (2011). "Perceptual consciousness overflows cognitive access". In: *Trends in Cognitive Sciences* 15, pp. 567–575.

— (2014). "Rich conscious perception outside focal attention". In: *Trends in Cognitive Sciences* 18, pp. 445–447.

Carruthers, P. (2000). *Phenomenal Consciousness*. Cambridge: Cambridge University Press.

Chalmers, D. J. (1996). *The Conscious Mind*. Oxford: Oxford University Press.

Cohen, M. A. and D. C. Dennett (2011). "Consciousness cannot be separated from function". In: *Trends in Cognitive Sciences* 15, pp. 358–364.

Dehaene, S., J.-P. Changeux, L. Naccache, J. Sackur and C. Sergent (2006). "Conscious, preconscious, and subliminal processing: A testable taxonomy". In: *Trends in Cognitive Sciences* 10, pp. 204–211.

Dennett, D. C. (1988). "Quining qualia". In: *Consciousness in Contemporary Science*. Ed. by A. J. Marcel and E. Bisiach. Oxford: Oxford University Press, pp. 42–77.

— (1991). *Consciousness Explained*. Boston, MA: Little, Brown & Company.

— (2005). *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge, MA: MIT Press.

Frances, B. (2008). "Live skeptical hypotheses". In: *The Oxford Handbook of Skepticism*. Ed. by J. Greco. Oxford: Oxford University Press, pp. 225–244.

Frankish, K. (2012). "Quining diet qualia". In: *Consciousness and Cognition* 21, pp. 667–676.

— (2016a). "Illusionism as a theory of consciousness". In: *Journal of Consciousness Studies* 23, pp. 11–39.

— (2016b). "Not disillusioned: Reply to commentators". In: *Journal of Consciousness Studies* 23, pp. 256–289.

Graziano, M. S. A. (2016). "Consciousness engineered". In: *Journal of Consciousness Studies* 23, pp. 98–115.

Hatfield, G. (2003). "Behaviourism and psychology". In: *Cambridge History of Philosophy, 1870–1945*. Ed. by T. Baldwin. Cambridge: Cambridge University Press, pp. 640–648.

Hull, C. L. (1943). *Principles of Behavior*. New York, NY: Appleton-Century.

Humphrey, G. (1951). *Thinking*. London: Methuen.

Humphrey, N. (1992). *A History of the Mind: Evolution and the Birth of Consciousness*. New York, NY: Simon and Schuster.

— (2011). *Soul Dust: The Magic of Consciousness*. Princeton, NJ: Princeton University Press.

Irvine, E. (2011). "Rich experience and sensory memory". In: *Philosophical Psychology* 24, pp. 159–176.

Irvine, E. (2012). *Consciousness as a Scientific Concept: A Philosophy of Science Perspective*. Dordrecht: Springer.

— (2013). "Measures of consciousness". In: *Philosophy Compass* 8, pp. 285–297.

Kind, A. (2001). "Qualia realism". In: *Philosophical Studies* 104, pp. 143–162.

Lamme, V. A. (2006). "Towards a true neural stance on consciousness". In: *Trends in Cognitive Sciences* 10, pp. 494–501.

Lau, H. (2008). "Are we studying consciousness yet?" In: *Frontiers of Consciousness*. Ed. by L. Weiskrantz and M. Davies. Oxford: Oxford University Press, pp. 245–258.

Levine, J. (2001). *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.

Mach, E. (1911). *The History and Root of the Principle of Conservation of Energy*. Chicago, IL: Open Court.

Mallon, R., E. Machery, S. Nichols and S. P. Stich (2009). "Against arguments from reference". In: *Philosophy and Phenomenological Research* 79, pp. 332–356.

Mandik, P. (2016). "Meta-illusionism and qualia quietism". In: *Journal of Consciousness Studies* 23, pp. 140–148.

Nida-Rümelin, M. (2016). "The illusion of illusionism". In: *Journal of Consciousness Studies* 23, pp. 160–171.

Pereboom, D. (2011). *Consciousness and the Prospects of Physicalism*. Oxford: Oxford University Press.

Phillips, I. (2011). "Perception and iconic memory". In: *Mind and Language* 26, pp. 381–411.

Prinz, J. (2016). "Against illusionism". In: *Journal of Consciousness Studies* 23, pp. 186–196.

Quine, W. V. O. (1980). "On what there is". In: *From a Logical Point of View*. Cambridge, MA: Harvard University Press, pp. 1–19.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.

Sandberg, K., B. Timmermans, M. Overgaard and A. Cleeremans (2010). "Measuring consciousness: is one measure better than the other?" In: *Consciousness and Cognition* 19, pp. 1069–1078.

Schwitzgebel, E. (2016). "Phenomenal consciousness, defined and defended as innocently as I can manage". In: *Journal of Consciousness Studies* 23, pp. 224–235.

Searle, J. R. (1997). *The Mystery of Consciousness*. London: Granta Books.

Skinner, B. F. (1953). *Science and Human Behavior*. New York, NY: Macmillan.

Sligte, I. G., A. R. Vandenbroucke, H. S. Scholte and V. A. Lamme (2010). "Detailed sensory memory, sloppy working memory". In: *Frontiers in Psychology* 1, p. 175.

Sloman, A. and R. Chrisley (2003). "Virtual machines and consciousness". In: *Journal of Consciousness Studies* 10, pp. 113–172.

Strawson, G. (1994). *Mental Reality*. Cambridge, MA: MIT Press.

Timmermans, B. and A. Cleeremans (2015). "How can we measure awareness? An overview of current methods". In: *Behavioural Methods in Consciousness Research*. Ed. by M. Overgaard. Oxford: Oxford University Press, pp. 21–46.

Titchener, E. B. (1899). *A Primer of Psychology*. New York, NY: Macmillan.

Tye, M. (2002). "Visual qualia and visual content revisited". In: *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press, pp. 447–456.

Watson, J. (1913). "Psychology as a behaviorist views it". In: *Psychological Review* 20, pp. 158–177.

Wittgenstein, L. (1958). *Philosophical Investigations*. 2nd ed. Oxford: Blackwell.