

David Hodgson

Goodbye To Qualia And All That?

Review Article

Max Bennett is a distinguished Australian neuroscientist, Peter Hacker an Oxford philosopher and leading authority on Wittgenstein. A book resulting from their collaboration, *Philosophical Foundations of Neuroscience*,¹ has received high praise. According to the Blackwell website, G.H. von Wright asserts that it ‘will certainly, for a long time to come, be the most important contribution to the mind–body problem that there is’; and Sir Anthony Kenny says it ‘shows that the claims made on behalf of cognitive science are ill-founded’.

The book builds on Wittgenstein’s remark that ‘Only of a human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees, is blind; hears, is deaf; is conscious or unconscious’ (quoted at p. 71). The authors identify what they call the *mereological fallacy*, the fallacy of attributing to a part of something properties that are correctly attributed only to the whole. Much of the book is a development of the claim that most neuroscientists commit this fallacy by attributing to brains properties and activities that can properly be attributed only to persons.

I won’t give a general review of the book, which does make valuable points concerning the importance of using language accurately in discussing mental concepts: helpful and laudatory reviews can be found on the Notre Dame Philosophical Reviews website (by Dennis Patterson) and in *Philosophy* 79, No. 307 (January 2004) 141–6 (by Daniel N. Robinson). However, I believe that some of its basic propositions are themselves fundamentally mistaken, and suggest that this is a consequence of disregard of opposing considerations, and insufficient recognition of the flexibility of language.

I will discuss three basic propositions from the book, which are particularly relevant from the ‘consciousness studies’ point of view. The first, my main target, is ‘to see an object is neither to see nor to construct an image of an object’ (p. 141). The other two are closely related: ‘It is mistaken to suppose that the subject of

Correspondence:

David Hodgson, Supreme Court of NSW, Queens Square, Sydney, NSW 2000, Australia.

[1] **M.R. Bennett & P.M.S. Hacker**, *Philosophical Foundations of Neuroscience* (Oxford: Blackwell, 2003).

experience has access or privileged access to his own experience' (p. 295); and 'The subject of experience is not ... a "self", but a human being' (p. 333).

(1) '*To see an object is neither to see nor to construct an image of an object*'

This assertion occurs towards the end of a discussion of the binding problem. Having argued that there is no binding problem as generally understood, because there is no sense in which features of a perceived object have to be 'combined in the brain', the authors continue:

Above all, to see an object is neither to *see* nor to *construct* an image of an object. The reason why the several neuronal groups must fire simultaneously when a person sees a coloured three-dimensional object is not because the brain has to build up a visual image or create an internal picture of objects in the visual field. When we see a tree, the brain does not have to (and could not) bind together the trunk, boughs and leaves, or the colour and the shape, or the shape and the movement of the tree. ... Which neuronal groups must simultaneously be active in order to achieve optimal vision, what form that activity may take, and how it is connected with other parts of the brain that are causally implicated in cognition, recognition and action, as well as in co-ordination of sight and movement, are what needs to be investigated by neuroscientists. Since seeing a tree is not seeing an internal picture of a tree, the brain does not have to construct any such picture. It merely has to be functioning normally so that we are able to see clearly and distinctly. It does not have to take a picture apart, since neither the visual scene nor the light array falling upon the retinae are pictures. It does not have to put a picture back together again, since what it enables us to do is to see a tree (not a picture of a tree) in the garden (not in the brain).

I agree that to see an object is not to *see* an internal picture of the object, but I contend it is beyond any doubt that we do see objects via images constructed by our brains.

I look at the setting sun just before it disappears below the horizon. In my visual field, there is something that is orange and apparently circular. What is it? The authors would say it is the sun itself, not an image of the sun.

However, if it is literally the sun itself, it is not the sun as it *is* when there is this something in my visual field, but as it *was* eight minutes earlier. Suppose the sun disintegrates during this eight-minute period. Then, if the orange something is literally the sun itself, rather than some kind of image or visual representation of the sun, the sun both exists (it is right there, right now) and does not exist (it has disintegrated) at the time when I am aware of the orange something. Further, the events or processes that are constitutive of there being this something in my visual field are much more closely linked to events or processes in my brain than to the most relevant events or processes on the sun: the latter events or processes took place 150 million kilometres from me and eight minutes earlier.

The orange something must surely be an image of the sun constructed by my brain (here meaning the whole mental and physical system of my brain and mind), though not in the sense of a *picture* of the sun that would itself have to be seen. Indeed, I do not *see* this image, and it is not any kind of thing that exists independently of being part of a conscious experience. It is unlike a picture in

that it is immediately present to me, and in that its existence is entirely dependent on my existence as a subject immediately acquainted with it. But it is an image nonetheless, in the sense that it is not the sun itself but rather is a visual representation to me of the sun. The authors rightly contend that, in normal usage, an image is something that we can recognise as an image because we can see it, and can compare it to the object of which it is an image and which we can also see. But this less usual use of the word is also correct.

Language developed initially in dealing in a common-sense way with everyday matters, but it is wrong to assume that language is limited to such uses. It can be applied to other than everyday matters in ways that go beyond common sense, perhaps with some use of technical words but also with much use of ordinary words, which then sometimes acquire different shades of meaning. This does not necessarily mean that these ordinary words have been given a special technical meaning, or even have become ambiguous. It is often simply a matter of the flexibility of language enabling new applications.

Suppose that prosthetic eyes are developed, and that the output from these eyes to the optic nerves can be artificially controlled by computational processes (not depending on light entering these eyes) so as to be indistinguishable from the output that results from actually looking with these eyes at the setting sun. The person using these eyes will then undoubtedly have a visual experience of the setting sun (of the nature of 'virtual reality'), through what can reasonably be called an image of the setting sun, constructed by the person's brain from the processes artificially induced in the person's optic nerves. And if so, it cannot be doubted that when the person actually looks at and sees the setting sun, so that similar processes are induced in the optic nerves by the light from the sun entering the eyes, the person's brain constructs a similar image of the sun and it is by means of that image that the person can see the sun.

From private communication with the authors, I believe that they would dismiss as mere illusion the case of an experience caused by artificially induced processes not depending on light entering the eyes. However, even if one calls this visual experience an illusion, it is still an illusion *of the setting sun*, and it is constructed from the artificially induced processes by a properly-working brain, operating similarly in this respect to the way it operates when the sun is actually seen. I see nothing problematic with the notion that we see things via images constructed by our brains and that our understanding of the appearance of things comes through these images; and this is by no means inconsistent with our appreciating that these images are of things that exist independently of us. Indeed, how could we know what things look like if the brain did not construct images of them from light coming from them?

To give another example, if I see a tree in my garden, it seems to be fixed in one spot, despite normal eye movements. But when I push with my finger against my right eyeball, and jiggle it, I seem to see two trees, one still stable as before and the other jiggling as I jiggle my eyeball. Before I jiggled my eyeball, the three-dimensional image of the tree constructed by my brain was stable because my brain had the ability, by complex information-processing, to compensate for

changes caused by my eye movements to the patterns of retinal nerve excitation. However, my brain does not compensate for changes to the patterns of nerve excitation caused by external jiggling of my eyeball, so the image present to my consciousness associated with the patterns of nerve excitation on the retina of my right eye can no longer be fused with that from my left eye into a single stable three-dimensional image; as a consequence there are now two two-dimensional images, one stable and one jiggling.

It does not merely *seem as if* there is something jiggling, there *is* something tree-like that is jiggling. What is the jiggling ‘tree’, if not an image of the tree constructed by my brain, based on the jiggling patterns of nerve excitation on the retina of my right eye?

I would add that this line of argument also challenges another basic assertion of the book, namely that there are no such things as qualia, understood as qualitative characteristics of experience. In the hypothetical case of an illusion of the sun induced by artificially produced outputs of prosthetic eyes, orange must be a qualitative characteristic of the experience (there is nothing else relevant that is orange); and I suggest that orange must also be a qualitative characteristic of the experience that occurs when similar processes of the optic nerves are produced by looking at the setting sun.

(2) *‘It is mistaken to suppose that the subject of experience has access or privileged access to his own experience’*

So long as it is denied that the brain constructs images of which we are immediately aware, this claim may seem plausible. But once it is accepted that the brain does construct such images, it becomes undeniable that the subject of experience alone is aware of these images; and this can reasonably be called privileged access to such images.

In aid of their claim, the authors (pp. 97–103) adopt Wittgenstein’s private language argument, to the effect that our use of language concerning experiences shows that the criteria for application of such language must be public. However, the regular (though not universal) association between various subjective experiences and observable events can explain how the meaning of words referring to subjective experiences can be learnt; and a person’s awareness of his or her own experiences is plainly of a different order from anyone else’s awareness of those experiences. It is true that the criteria for the application of psychological concepts such as pain cannot be entirely private; but we come to know that in relation to concepts like pain it is the private criteria that are the most critical for the correct use of the word.

(3) *‘The subject of experience is not ... a “self”, but a human being’*

If it is accepted that the brain constructs images which are immediately present in consciousness, and to which the subject has a kind of privileged access, then questions arise as to the nature of the subject.

It is true that, in ordinary use of language concerning subjects of experience and subjects of action (i.e. agents), it is generally a person as a whole (or animal as a whole) that is taken to be the subject. We do not say that a person's brain has this or that experience, or does this or that, but that the person has this or that experience, or does this or that. However, even in ordinary use this is not always the case. People do ask what will happen to them after they die, and this question is understood as relating to them as subjects of experience and action, and is not dismissed as nonsense. The great religions make assertions about an afterlife, or about reincarnation, and these assertions are widely discussed and disputed rather than being summarily dismissed as nonsense.

Further, although in ordinary language the subject is taken to be a person as a whole, not all of a person is necessary to constitute a subject. I see a tree. I would still see it if I lost my arms and legs. I would still see it if the life-support given to my head and brain by the rest of my body were to be provided artificially. I would still have a similar experience if the nerve processes from my eyes to my brain were to be provided artificially. I think it is reasonable therefore to say that, on close analysis, the subject of this experience can be considered as being something other than the whole human being, and there can be meaningful consideration of what that something is.

On the other hand, living persons can entirely lose the capacity to be a subject, for example in a coma. So even living persons are not necessarily subjects, and subjects are not necessarily complete living persons. What is essential about subjects is that they have the capacity to experience and/or act. My contention is that, while we normally equate subjects with persons (or animals) as wholes, we combine this with a reservation of judgment as to what is necessary and sufficient to constitute a subject. And I contend this is largely an empirical question.

From private communication with the authors, I believe they would say that there is a prior conceptual question, namely what would satisfy behavioural criteria that would warrant ascribing to it the relevant psychological predicates. However, it follows from my earlier discussion that the correct ascription of at least some psychological predicates depends crucially on what if anything is experienced, not on behaviour. And I contend that it is conceivable that experiences could occur even if there is no person or animal having them, for example in an artificially created system or in an afterlife; and that in any such case there would be a subject of experience. While we are clearly a very long way from creating an artificial system that could have an experience, and while I do not believe there is an afterlife in the sense of some kind of straightforward survival of death, the questions whether such a system could be created and whether there is an afterlife, in either case with an associated subject of experience, are not resolved by mere conceptual arguments.

I would add, in conclusion, that like Bennett and Hacker I support a non-reductionist approach to questions concerning the brain and mind; but I contend that this approach needs to be supported by empirical as well as conceptual considerations.