

Do Brains Think?

Christopher Humphries, London, England, UK

1 Introduction

The motivating idea of B&H's 2003 *Philosophical Foundations of Neuroscience* ('PFN') is that a clear view of the relationship between neuroscience and human psychology is not possible without a correct analysis of the psychological concepts and categories involved in the descriptive understanding of mental life. The authors find that these concepts are often misconstrued or misapplied by neuroscientists and their philosophical allies. Defective understanding and misguided questions may, at worst, render research futile by misdirection of experimentation and misunderstanding of its results. It is the authors' constructive intention that their conceptual analysis should 'assist neuroscientists in their reflections antecedent to the design of experiments.'

A leitmotif of PFN is the identification of a persistent mistake of construing the brain, or components of the brain, as subject or locus of mental predicates. For B&H, the ascriptions properly belong to the person or animal. The mistake institutes a sort of Cartesian revanchism, with the old error of ascribing psychological attributes to a mental substance replaced, in the new materialist version, with the error of ascribing them to a physical substance. Brain/body dualism is incoherent, like talk of the East Pole. Thus (PFN: 71): 'Only of a human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees, is blind, is deaf; is conscious or unconscious.' (Wittgenstein 2001: I §281); and 'Perhaps indeed it would be better not to say that the soul pities or learns or thinks but that the man does *in virtue of the soul*.' (Aristotle 1986: 408b).

The neuroscientist's reply might be that talk of brains and their neural circuits seeing shells flying and deciding to take cover is an innocent *façon de parler*; a harmless and amusing shorthand that leads to no practical error. Its value is metaphorical: for example, when describing neural mechanisms, it can harness the insights that have accrued to neuroscience from the field of information technology. For B&H, this last is further confusion: brains are not computers, and computers do not enact rule-governed manipulation of symbols. Computers are artefacts that 'produce results that will coincide with rule-governed, *correct* manipulation of symbols.' (Bennett and Hacker 2007: 151). The projection of the designer's perspective into the operation of the computer is a version of the very error of thought currently in view.

B&H assert a sharp line between investigation of the logical relations between concepts – the philosopher's trade, having to do with the distinction of sense and nonsense – and the scientist's investigations, which have to do with empirical truth and falsehood. But the orthogonality of truth and sense is assailable: e.g. are not answers to conceptual questions true or false? (Dennett 2007: 79-82) Again, B&H's claim that conceptual truths delineate the logical space in which the facts are located, and are prior to them, (129) could be met by the simple objection that the concept of colour is not prior to colour facts (cf. PFN: 129-130). At the opposite pole to B&H is the Quinean view. Abandonment of the 'two dogmas of empiricism' results in a 'blurring of the supposed distinction

between speculative metaphysics and natural science.' Thus it is nonsense 'to speak of a linguistic component and a factual component in the truth of any individual statement.' Conceptual scheme and the deliverances of sense interpenetrate within our 'total science'. (Quine 1961)

2. An Inner Process

According to B&H, it only makes sense to ascribe mental predicates to what is or resembles a living human being. Following Wittgenstein, behavior is taken to provide logical criteria for the application of mental concepts. Only the person (the rational, responsible being), and not the brain, satisfies these criteria (PFN: 83). Searle takes this Wittgensteinian move to be at the heart of the argument that leads to the impossibility, for B&H, of consciousness, qualia, feelings etc. existing in and being predicable of brains (Searle 2007: 102). Further, Searle takes B&H to identify pain (let's say) with the criterial basis for pain, i.e. its external manifestation. Then, because the pain is seen to be identified with its criterial manifestation, Searle takes B&H to think that it cannot be the subject of neurological investigation. On this understanding, the PFN programme amounts to criterial behaviourism:

Just as the old-time behaviourists confused the behavioral evidence for mental states with the existence of the mental states themselves, so the Wittgensteinians make a more subtle, but still fundamentally similar, mistake when they confuse the criterial basis for the application of the mental concepts with the mental states themselves. That is, they confuse the behavioral criteria for the *ascription* of psychological predicates with the *facts ascribed by these* psychological predicates, and that is a very deep mistake (103)... The fallacy, in short, is one of confusing the rules for using the words with the ontology (104)... I think that once this basic fallacy is removed, then the central argument of the book collapses. (105)

I don't think this charge sheet will hold up in court. In the first place, B&H nowhere explicitly make the equation between behavior and the subject ontology of mental predicates. The former is criterial for the latter, not identical with it. Pain behavior is a manifestation of pain, and a criterion of it, but is not the pain itself. Moreover, the charge of behaviourism is refuted if the behavioral criterion is 'defeasible', i.e. only partly constitutive of its object. Thus, if I'm reciting the alphabet 'in my head', there *is* no behavior. B&H display the defeasibility of behavior when they say: 'an animal may be in pain and not show it or exhibit pain behavior without being in pain. (We are no behaviourists.)' (Bennett and Hacker 2007: Note 18 p211).¹ Secondly,

¹ Wittgenstein takes behavior as criterial for the mental, but not to be equated with it ontologically or causally: the relation is logical and normative. Thus behavior, expressed by the body, is the window of the soul. (Wittgenstein 2001: II §178) Only to a being that has capacities can mental concepts be ascribed. But a being that has capacities can exercise them or not: the matter is not causally determined. Behaviourism is therefore no apt theory of such a being. See Glock: 55-58 and Hacker 1990: 224-254. Thus Wittgenstein is not a metaphysical behaviourist. Logical behaviourism (asserting semantic equiva-

B&H do not deny that it is possible to mount neurological investigations of pain etc. 'Research on the neurobiology of vision is research into the neural structures that are causally necessary for an animal to be able to see and into the specific processes involved in its seeing.' (Bennett and Hacker 2007: 161).

This last point raises the question of the causal relationship between pain etc. and neurophysiology. Searle acknowledges that B&H look for causally necessary conditions for consciousness, but insists that a causally sufficient account is what is required, and uses this assumption in the construction of his case for B&H's Wittgensteinian behaviourism. But that requirement seems to beg the question about the dichotomy of mental and neurophysiological predicates, with a tacit assumption that a correct theory of mind must be physically reductive. For reduction requires explanatory connection between explanandum and explanans, together with bridge laws connecting the relevant properties. A causally sufficient explanation of consciousness in terms of physical law would deliver both of these requirements for reduction straight away. Reduction, thus established, would dissolve the possibility of a division of categories between the inner and the outer. So the hidden reductive assumption covertly imports the conclusion into the premises.

Searle's reply to this might follow his chapter essay 'Reduction and the Irreducibility of Consciousness' (Searle 1992: Chapter 5). Consciousness is there described as a 'causally emergent property of systems' on a notion of emergence that denies that an emergent has any causal powers that cannot be explained by the causal powers of the physical base. Searle says that this type of emergence usually delivers causal-explanatory reduction, from which ontological reduction follows. However, Searle continues, in the case of consciousness, the ontological reduction does not work, because the subjectivity of experience cannot be explained in third-party causal terms. But this failure of ontological reduction he claims to be unimportant, because it is a trivial consequence of our definitional practices. We cannot, following the usual reductive procedure, redefine consciousness in causal terms which, being causal, discount the appearances that are characteristic of the reduced domain, because in this case the appearances are what are of interest. On this argument, mental predicates related to consciousness are not ontologically reduced, and so the question is not begged.

However, even if this argument is accepted, it still does not go far enough. This is because, although it says why reduction is harmless in case of consciousness, it does not show why reduction would be harmless in case of normativity. B&H, as already noticed in the discussion of information technology, take normativity to be external to the causal realm as exemplified by the computing artefact. So Searle's critique of B&H's psychological ontology still begs the question for mental predicates related to normativity.

In ascribing mental predicates to the animal rather than the brain, B&H are proclaiming that the predicates, together with their ontological subjects, belong to a separate and distinct logical category. Searle criticises B&H's expression 'mereological fallacy', pointing out that brains are not proper parts of persons: what B&H are attacking is a would-be Rylean category mistake. Precisely

so. Most of the category mistakes on the table in PFN are simple logical mispredications, not requiring a specifically Wittgensteinian unmasking.

For Searle, it is 'more or less educated scientific common sense' that conscious states 'exist in the brain', being produced causally as 'higher-level or system features' (2007: 99). For B&H, neither does a mental predicate attach to a brain as subject or agent, nor is the mental fact referred to located *in* the brain. Thoughts do not occur in the brain, they occur in the study (PFN: 179-180). The claim that to deny the brain-location of thinking is like denying the stomach-location of digesting (Searle 2007: 109) exemplifies the tacit reductivism already noticed.

3. Persons

For B&H, the proper subject of mental predicates is the person, though no extended analysis is offered of what a person is. For that desideratum we may borrow a page from the patriarchs (Strawson 1957). Thus material objects are found to be the basic particulars – identifiable and re-identifiable without reference to other particulars and partly constitutive of the 'uniquely pervasive and comprehensive' system of individuation provided by time and space. Persons are a separate and distinct class of particulars, ascribing to themselves actions, intentions, thoughts, feelings, perceptions etc. These are predicated of a single entity, which is grammatically and, by argument, ontologically the same entity as that to which are ascribed the physical characteristics of the person. ('I am happy; I am thin.') This entity, the person, is logically prior to the individual consciousness, for if the priority is taken the other way round, no experience can ever be attributed other than to oneself. The ontological priority of the person must be accepted, not to avoid scepticism, but 'in order to explain the existence of the conceptual scheme in terms of which the sceptical problem is stated.' (Strawson 1957: 106) (Hacker rejects this 'dichotomous division of predicates' as 'overly Cartesian' and prefers a more vague definition of the person as a subcategory of the animal, having capacities of reason, will and morality; see Bennett and Hacker 2007: 312-3.)

A person, then, is subject of both physical and mental predicates. The Strawsonian analysis upholds the division of category between physical and mental predicates, while uniting them in the person. B&H's central point about mispredication is not a uniquely Wittgensteinian insight, but flows from a distinction of categories that is fundamental in the descriptive metaphysics of mind and body. The point does not therefore stand or fall with the various peculiarities of PFN, such as the claim that two people can share the same pain (in the way that two pillar boxes can share the same colour, see PFN: §3.8) and that subjective qualities of consciousness (qualia) do not exist (qualia not being properties of consciousness but of objects: 'quale' equivocates between the subjective quality of an experience and the experience itself, see PFN: §10.3).

lence of mental predicates and behavioral dispositions) is a stronger tendency, though less so in Wittgenstein's later thought.

4. Conclusion

So do brains think or don't they? B&H think not, and I have argued that their conclusion does not depend on their specifically Wittgensteinian account of contemporary neuroscience. The proposition can be denied, as a category mistake, from an alternative descriptive-metaphysical approach.

I think B&H's arguments are stronger than Searle's critique of them. But the Quinean point made above disrupts the neat conceptual taxonomy. The way in which scientific knowledge influences the *a priori* conceptual scheme is a large question, that cannot be analysed here. But this work is needed, because if the conceptual and the empirical are orthogonal in the way that B&H claim, then there is nothing further to be said about the ontology of mind: enquiry is brought to a close by their strictures.

Literature

- ARISTOTLE 1986 *De Anima*, tr. Lawson-Tancred, H., London: Penguin Books.
- BENNETT, MAXWELL, DENNETT, DANIEL, HACKER, PETER AND SEARLE, JOHN 2007 *Neuroscience and Philosophy: Body, Mind and Language*, New York: Columbia University Press.
- BENNETT, M.R. AND HACKER, P.M.S. 2003 *Philosophical Foundations of Neuroscience*, London: Blackwell.
- DENNETT, DANIEL 2007 "Philosophy as Naïve Anthropology: Comment on Bennett and Hacker" in: Bennett et al. 2007.
- GLOCK, HANS-JOHANN 1996, *A Wittgenstein Dictionary*, London: Blackwell.
- HACKER, P.M.S. 2007 *Human Nature: The Categorical Perspective*, London: Blackwell.
- HACKER, P.M.S. 1990 *Wittgenstein: Meaning and Mind*, London: Blackwell.
- QUINE, WILLARD VAN ORMAN 1961, "Two dogmas of empiricism", in *From a Logical Point of View: Nine Logico-Philosophical Essays*, 2nd ed., Cambridge Mass: Harvard University Press.
- SEARLE, JOHN 2007 "Putting Consciousness Back in the Brain: Reply to Bennett and Hacker" in: Bennett et al. 2007.
- SEARLE, JOHN 1992 *The Rediscovery of the Mind*, Cambridge Mass: The MIT Press.
- STRAWSON, PETER 1957 *Individuals: An Essay in Descriptive Metaphysics*, London: Routledge.
- WITTGENSTEIN, LUDWIG 2001 *Philosophical Investigations*, tr. Anscombe, G.E.M., London: Blackwell.