

EDITORIAL

Neurophilosophy¹

Neurophilosophy is Patricia Churchland's (1986) hypothesis that the mind–brain 'problem' and related 'problems' such as those surrounding perception and cognition will be solved by advances in cognitive neuroscience. This entails that the attempts by philosophers to solve these problems by their methods of purely logical, conceptual and linguistic analysis are doomed to failure as they represent a mistaken strategy. Today is certainly an exciting time in cognitive neuroscience when we are witnessing a continual stream of important discoveries on how the brain works. The dominant themes include parallel distributed processing (PDPs), computing and information processing by nerve nets (both real and computer-simulated), Hebbian synapses, neuronal cell assemblies, the role of chaos and non-linear dynamics, unit recordings in awake behaving animals, PET scanning of ongoing human brain function in specific tasks and a wealth of basic information concerning the neurochemical and biophysical basis of learning, perception, memory, etc.

Churchland also presents two other planks in her platform. The first is the belief that the Identity Theory (IT) of mind–brain relation is correct, or will be shown to be so by advances in neuroscience. The second is the recommendation that 'folk' ways of talking about mind, seeing, thinking, feeling, knowing, etc. will gradually be replaced by the way that neuroscience talks about these functions.

I am a firm supporter of the position that what are currently widely regarded as 'philosophical' problems about how the mind relates to its brain will eventually be solved by science and not by conceptual or linguistic analysis. But I also believe that neurophilosophy, at present, has too limited a concept of what sciences are relevant. It somewhat neglects introspection science and it assumes that contemporary physics is to all intents and purposes complete. Furthermore, it shares in the hopeless confusion that afflicts some of our very basic scientific concepts in this area.

To deal with the latter first.

(1) No IT theorist ever applies the theory to his or her own consciousness. For this they rely firmly on the 'folk' theory of naive or Direct Realism (DR), however much in theory they support the rival scientific theory of perception known as the representative theory (RT). At the base of this schizophrenia is the (almost) universal confusion between the 'body image' of neurology (i.e. the somatic sensory field) with the physical body itself (see Schilder, 1950; Smythies, 1953 for details).

(2) This is compounded by a second (almost) universal confusion between the visual field and the stimulus field. The visual field (VF) is located in consciousness and is filled with our visual sensations – the product of the vastly complex representative mechanisms of perception. We can only observe these latter as phenomena by introspection, as in the many investigations by introspection psychologists such as Klüver, Gregory, Vernon, myself and the whole Gestalt school. The stimulus field (SF), in contrast, consists of the physical objects in the external world that my visual field represents. The SF lies at the beginning of the complex causal chain of perception. We examine the SF field by exteroception (perceiving). We examine the VF by introspection (sensing). A good model for this is the TV screen where we can examine it as it is in itself made up of a raster of parallel coloured lines: or we can use it to observe the events in the TV studio. Since many neuroscientists call the SF the VF confusion abounds. A further complexity is provided by the many retinotopic maps in visual cortex that contain coded information about the SF.

(3) Then there is much confusion about the use of the terms 'conscious' and 'consciousness'. We must distinguish between at least three quite different meanings: (i) the medico-behavioral sense, e.g. 'the patient is conscious now, doctor'; (ii) the attentional sense, e.g. 'she was conscious of an

¹ Dr J. R. Smythies, Department of Neuropsychiatry, Institute of Neurology, Queen Square, London WC1.

erie feeling in the room'; and (iii) the phenomenal sense in which we define phenomenal consciousness as the total of all that we can introspect – our sensory and image fields, feelings, thoughts and the Self ('Ego' – the 'O' of introspection psychology).

(4) Fourthly there is the general confusion between phenomenal space and physical space. The fault here lies with Descartes who made the crucial error in supposing that all mental entities lack extension in space. Rorty (1970) makes the key distinction between mental occurrents (such as pains and sense impressions) and mental features (such as beliefs, desires, hopes, etc.). Now, once we have realized that DR is incompatible with a scientific account of perception, it becomes clear that some mental particulars at least i.e. visual and somatic sensations and images – are extended in space. If you examine one of your own after-images, for example, you will clearly see that this is so. The after-image floats before you located in your visual field. It is clearly extended in space. It has a spatial boundary and topologically divides the whole of your visual field into two parts – one inside and one outside – thus, it is an instance of a topological entity known as a Jordan curve. Of course, the after-image is not located 'out there' in physical space, neither is your visual field. 'Out there' in this context is relative to 'O' and particularly to your body-image. It is not 'out there' relative to your physical body. All phenomenal events, including after-images, are located in internal phenomenal space and not in external physical space. So the problem can be reduced to the question 'How is phenomenal space related geometrically and topologically to physical space?' which I have tackled elsewhere (Smythies, 1992). The Cartesian error that no mental elements can be extended has crippled progress in this field for three centuries. It arose because he confused the Ego and its thoughts with the totality of consciousness.

To return to IT: this states that mental events M are identical to certain brain events B . This leads to no problems so long as we are dealing with Rorty's mental features. But when we turn to deal with his mental occurrents, difficulties abound.

(i) The brain contains only extremely complex patterns of electrical potentials located in its nerve nets. These patterns, and their interactions, code a wide number of features of the environment. The neurons process this information and all this activity impinges on the motor cortex so that coherent behaviour results adapted to the demands of the environment. But the brain has a second 'output' besides motor behaviour, and that is conscious experience, in particular the phenomena of our sensory fields. Lord Brain (1960) has pointed out that the latter are uncoded representations of the events in the stimulus field. So how does the brain actually construct, in engineering terms, the visual field? This is Brain's 'decoding' problem that has recently resurfaced in neuroscience as the 'binding' problem. If we take TV as a model for certain aspects of perception (Grey-Walter, 1950; Charlesworth, 1979), this precise function is done by the raster mechanism of the set. What then could be the analogue of the raster in the brain?

(ii) Churchland (1986) claims that obtaining more information about B will eventually demonstrate beyond all reasonable doubt that $B = M$ (we must take M here as introspectable mental occurrents). But we also have a lot of information about M obtained by introspection psychologists, and, if $B = M$, then of course $M = B$. So information about M will, or should, give us information about B (of which M are supposed to be working parts). But there are no structures in B that remotely resemble the structures of M . That is, information contained in the VF in M is information about the stimulus field, and this information is coded in B , but uncoded in M . And it is axiomatic that x and a coded version (x') of x cannot be identical. x' represents x but cannot be identical to x .

As Gunner (1967) says, Churchland's claim (and Smart's before her) is untenable. To demonstrate that $M = B$, it is not enough to pile up information about M , nor about B , as Ayer (1950) pointed out long ago. You have to demonstrate by scientific means (after all IT claims to be a scientific theory (Gregory, 1984)), that M in fact = B . Or at least you should be able to suggest some line of investigation that might lead eventually to such a demonstration. Recently two series of experiments have produced results that some people have claimed do just that.

Salzman *et al.* (1990) carried out microelectric stimulation of small groups of neurons in the columns of area MT in the awake, behaving monkey. These cells had visual directional properties.

The experimenters then showed that the monkey's behaviour, which had been previously conditioned to such an *S*, was such that it indicated that the monkey had actually seen that *S*, the visual movement for which these cells were selectively tuned. Albright (1992) concludes from this data that '...perceptual state is clearly coupled to discrete brain events further strengthening our belief in their identity'. This conclusion does not, however, follow. The same argument can be used to prove that percepts are located in the retina, if I microstimulate cells there. All these microstimuli merely result in hallucinations, and any hallucinatory *S* can perfectly well trigger behaviour that has previously been conditioned to a veridical version of that *S*. Of course no one denies that these cells mediate the perception of movement. But that is not the point at issue, which is: 'Are the percepts (sensations) identical with the activity of these neurons?' The counter claim that can always be made is that the actual percepts are located 'higher up' in the system.

In a second series of similar experiments (Logothetis & Schall, 1989), unit recordings were taken from single neurons in the MT area of awake, behaving monkeys. The *S* was one of two horizontal grids, one moving up and the other moving down, presented to each eye separately so that retinal rivalry resulted. The monkey was trained to execute an *L* saccade in response to one *S* and an *R* saccade in response to the other. It was then shown that some of the neurons were signalling the *S* present in that neuron's receptive field, whereas other neurons were signalling what the monkey was seeing (i.e. the pattern in the dominant eye), as judged by the direction of the saccade. An explanation for this in terms of pure motor functions was ruled out. Can we assume, as some have done, that this experiment demonstrates that $M = B$ in this case? Logothetis & Schall themselves do not as they say that it cannot be ruled out that '... the perception related modulation observed in the neuron may be the result of feedback from higher centers'. This derives from the infuriating logic of all representative mechanisms, which is that one can never determine by looking only at one site in the mechanism where the phenomena observed actually originated. That is why TV stations sometimes flash the signal on your screen 'Do not adjust your set'.

The only advantage of IT is that its only rivals (if we exclude fluff like functionalism), such as Cartesian dualism, are in even worse shape. I would put my money on the suggestion made by Penrose (1989), Linde (1990) and Chomsky (1975) that physics itself will have to undergo a revolutionary expansion of some kind before we will be able to understand the mind and its relation to the brain (Smythies, 1992).

J. R. SMYTHIES

REFERENCES

- Albright, T. D. (1992). Motion perception and the mind-body problem. *Current Biology* (in the press).
- Ayer, A. J. (1950). Comments in *The Physical Basis of Mind* (ed. P. Laslett), pp. 65–66. Blackwell: Oxford.
- Brain, Lord (1960). Space and sense-data. *British Journal for the Philosophy of Science* 11, 177–191.
- Charlesworth, M. (1979). Sense-impressions: a new model. *Mind* 88, 24–44.
- Chomsky, N. (1975). *Reflection on Language*, p. 24. Random House: New York.
- Churchland, P. S. (1986). *Neurophilosophy*. MIT Press: Cambridge.
- Gregory, R. L. (1984). *Mind in Science*. Penguin: Harmondsworth.
- Grey-Walter, W. (1950) Features in the electrophysiology of mental mechanisms. In *Perspectives in Neuropsychiatry* (ed. D. Richter), p. 77. H. K. Lewis: London.
- Gunner, D. L. (1967). Professor Smart's 'Sensations and Brain Processes'. In *The Identity Theory of Mind* (ed. C. F. Presley), pp. 1–20. University of Queensland Press: St Lucia, Brisbane.
- Linde, A. (1990). *Particle Physics and Inflationary Cosmology*. Harwood: New York.
- Logothetis, N. K. & Schall, J. D. (1989). Neural correlates of subjective perception. *Science* 245, 761–767.
- Penrose, R. (1989). *The Emperor's New Mind*. Oxford University Press: Oxford.
- Rorty, R. (1970). In corrigibility as a mark of the mental. *Journal of Philosophy*, 77, 399–424.
- Salzman, C. D., Britten, K. H. & Newsome, W. T. (1990). Cortical stimulation influences perceptual judgements of motion direction. *Nature* 346, 174–177.
- Schilder, P. (1950). *The Image and Appearance of the Human Body*. International Universities Press: New York.
- Smythies, J. R. (1953). The experience and description of the human body. *Brain* 76, 132–145.
- Smythies, J. R. (1992). *The Walls of Plato's Cave*. (In the press.)