

## TONY CHADWICK ESSAY PRIZE 1997 WINNER

### PHILOSOPHY OF MIND - A 50s VIEW OF THE 90s

by Alan Bailey

It may be of interest to set down some reflections on current discussions of the 'mind-body problem', and in particular Michael Lockwood's classes on 'The Enigma of Consciousness', from the perspective of an interval of 40 years since last being involved with philosophy at Oxford.

It is obvious that the whole way of doing philosophy has changed radically. No more talk of language ('what we would say...') and instead a great many empirical statements, based largely on recent work in the relevant sciences (physics, neurophysiology, evolutionary biology, computer science).

This has some clear advantages. If the object is to provide a (more) satisfactory conceptualisation of experience, the current 'science-based' approach gives a much richer range of concepts than 50s 'ordinary language' - as Michael Lockwood put it, 'I don't think our usual categories are very much help - the concepts of physics are much more fruitful'.

Another advantage is that it allows, indeed encourages, speculation about large-scale propositions and theories which are beyond the established limits of current science, but where it is possible to say what evidence would in principle count as validating them (or not), and where it may be possible at some future time actually to get the evidence to establish them (or not). Even in the 50s some of us (a minority) thought that this was a way forward in philosophy, allowing large-scale theoretical speculation, testable only in principle, in areas such as the philosophy of history, philosophy of mind, philosophy of religion or philosophy of science (though most of us knew very little physics).

But this was seen as very dangerous ground. The influence of the Viennese logical positivists and Wittgenstein was very strong. If such speculations could not be tied down by testable evidence, in the way that physical theories could be 'tied down' (not just crudely verified or falsified, but at least found to fit the facts better or worse), then they were seen as making no sense and not worth uttering. From this point of view I still have difficulty with some applications of the Nagel criterion of 'knowing what it is like to be...' It works well for the inference to other minds - from our own introspective experience to what it must be like for other (normal, thinking) people, as the best explanation of their observed behaviour. Here it is not just an imaginative leap into the unknowable, but has a 'cash value' (as the 50s would have said) in explaining the behaviour - in a way in which we would not want or need to explain the 'behaviour' (i.e. output) of a computer. It is good to find that it is now again apparently respectable to accept the evidence of introspection, as against behaviourism (and, perhaps, Ryle's position in *Concept of Mind*, though that was never clear). But ascribing consciousness on this criterion to lower orders of animate creatures seems to me problematical. Though we can imagine what it is like to be a gnat or a prawn, it is not merely untestable (as indeed it is for other minds, in any direct sense) but there is no element in their behaviour for which this postulation of consciousness is clearly the best explanation. Nagel asserts heatedly that there is nothing wrong with this: 'To deny the reality or logical significance of what we can never describe or understand (e.g. experiences of bats or Martians) is the crudest form of cognitive dissonance' -

but of course the question is whether we can ever make any sense of such speculation and therefore whether it is worth discussing.

There are perhaps two candidates for that extra element in behaviour which we would require for mental language to be appropriate (in 90s language, for the hypothesis of consciousness to be the best explanation):

(i) Behaviour which is unpredictable and therefore apparently random - the kind of 'behaviour' which makes us talk about machines (cars, computers) in 'personal' terms ('temperamental', 'behaving badly'). But this is of course metaphorical talk, because in those cases we expect there to be some mechanical explanation such that the behaviour would be predictable if we knew more of the physical facts. And in any case, treating random behaviour as a necessary attribute of people has its own problems, and seems to have little to do with consciousness (though there could be a connection, given the view of consciousness as a detached 'corrective' to automatic behaviour).

(ii) The sub-set of behaviour which counts as 'communication'. In human terms this generally involves language; thus a conclusive test to show that someone paralysed is conscious is if some way can be found of enabling them to give (verbal) answers to questions. But non-verbal communication would do; in animals such as domestic pets, behaviour such as tail-wagging or simply 'the look in his eye' encourages us to say 'He knows/understands' and hence '...is conscious'. But this depends on our inference from this humanlike behaviour that 'what it is like to be a dog' is sufficiently like what it is (introspectively) like to be us. Communication between creatures of the same species (chimpanzees, bees) does not by itself warrant the same inference.

The 50s anti-metaphysical stance is also hard to square with the Kant/Russell/Lockwood position on 'things-in-themselves'. It is true that the account of reality given by physics (as opposed to common sense) turns out to be a mathematical, abstract description, which can only be related 'at its edges' to observable phenomena. That is the nature of theories (best explanations), and the grander the theory the more remote its propositions tend to be from empirical observation, let alone ordinary experience. But the two moves that are made going beyond this position seem to launch into metaphysics:

- All knowledge of the physical world is purely 'structural' requiring some different means of access to know what things are 'in themselves'.

- When we are looking at something we are conscious of physical qualia (in the brain) which provide 'a glimpse of the intrinsic nature of the physical world' (as 'eigenvalue spectra of observables').

The first of these moves seems to me false (inconsistent with the earlier view about empirical input 'at the edges') and the second unintelligible (in the 50s sense of 'untestable' without 'cash value'). Of course the non-scientific 50s view has great difficulty in understanding this in any detail, but that does not necessarily invalidate its scepticism.

Another striking feature of the 90s is that scientific 'explanation' has been elevated into the sole criterion of what is 'real' as opposed to direct observation. This is stated explicitly in Paul Checkland's *Matter and Consciousness*: 'If the new framework is far better than the old at

explaining and predicting phenomena, then we have excellent reason for believing that the theoretical terms of the new framework are the terms which describe reality correctly' (my underlining). Hence he says flatly: 'The red surface of an apple does not look like a matrix of molecules reflecting photons at certain critical wavelengths, but that is what it is'. One would like to say that it is also, and equally, the red surface of an apple, but this seems to be ruled out as 'incorrect'. *A fortiori*, given these ground-rules, mind-events are 'really' brain-events. 'Current theories of meaning tend to reverse the positivist's view entirely: the meaning of any term, including observation terms, is fixed by its place in the network of beliefs in which it figures...Our mentalistic vocabulary, therefore, will have to be rejected, if at all (and why not?), on grounds of its explanatory and predictive shortcomings'. This seems to be eliminative materialism by diktat - physics *Uber alles*. And it leads to outrageous statements such as: 'If machines do come to simulate all of our internal cognitive activities, to the last computational detail, to deny them the status of genuine (presumably conscious) persons would be nothing short of racism'. As John Searle pointed out, a computer could simulate a hurricane without there being a hurricane in the computer. To discuss the mind-body problem without recognising introspection is like running a race without crossing the start-line.

One practical difficulty in deciding between identity theory and some form of dualism is that neurophysiological knowledge about 'supervenience' (the set of brain-events which invariably and simultaneously accompany a mind-event - Michael Lockwood's 'psycho-physical lexicon') is so sparse. So it may be worth trying a 'thought-experiment': suppose that someone has electrodes implanted in his visual cortex, so that whenever they are switched on (from some remote source) the colours that he sees appear to shift towards the red end of the spectrum but he is aware of no other sensation, and when the electrodes are turned off his vision returns to normal. At first, when he is asked to describe his sensations and the electrodes are switched on, he would say something like 'Everything suddenly looks redder'; but after a time, when he is used to the sensation and understands its cause, he would be more likely to answer a question about what has happened by saying 'The electrodes have been switched on' - on the general principle that it is unhelpful and even misleading to describe only sensations when something more 'objective' can be confidently inferred from this (e.g. it would be annoying for someone to say 'I can see a white hazy patch above the house', and only when pressed 'It looks like smoke!').

So in the experiment, 'Everything suddenly looks redder' and 'The electrodes have been switched on' can be seen as alternative descriptions of the same event. In terms of explanation, one description could be said to 'reduce to' the other. But from the subject's point of view there is no reason why he should accept that the former description is 'incorrect', or that what he is 'really' seeing is the electrodes being switched on (indeed the latter is plainly false). It still seems to me a matter for decision whether to count what has happened as one event or two. Of course more newly discovered facts could be relevant to that decision. But given what we already know, there is nothing 'objective' or 'intrinsic' in the situation to resolve the apparent difference between 'dualism' and 'identity theory'.

One way of dealing with this would be to adopt a version of the identity theory with such events being seen as having 'dual aspects' (a view mentioned but not I think adopted by Michael Lockwood). But of course this view is not without its difficulties, and this essay concludes by commenting on two of these:

(i) Intentionality: John Searle phrased the problem in terms of mental states having 'intrinsic semantics'. But does this perhaps apply only to a sub-set of mental states which we have learned to distinguish into types and talk about? - in other words, could 'intentionality' be an intrinsic feature of mental language, not of mental states themselves? This kind of explanation in terms of how we learn to use words may be unfashionable in the 90s, but if we start from introspection the capacity to describe and understand descriptions of what is going on in someone's mind certainly requires some explanation. On this view, poetry can be seen as a way of evoking mental states (moods, fleeting perceptions) without describing in standard prosaic terms what they are 'about'.

(ii) Unity: If the problem here were only Descartes' observation that the mind is not divisible in the way that all physical objects are divisible, then Michael Lockwood's suggestion that its physical counterpart could be a 'phonon single collective mode' with the same quality of indivisibility, could meet the case. But there may be a deeper problem, related to the essential 'subject-object' structure of mental states and events which he also discussed. Mental states/events do not simply lie around waiting to be knitted together by short-term memories etc. into recognisable 'personalities'. We do not know what it would be like for there to be a mental state/event which was not someone's; and on the other hand we cannot be aware of ourselves as conscious beings without being conscious of something. (Michael Lockwood's distinction between consciousness and the content of consciousness may be relevant here; but it is difficult to see what consciousness would be like apart from content - watching a blank TV screen with no aerial is not a particularly helpful analogy if we are imagining someone watching.) Clearly Descartes' *Cogito* was focusing on this essential 'subject-object' structure of mental events, hence is not equivalent to saying merely 'There are mental events (lying around)'. On this view consciousness is not just a process (Searle, c.f. digestion) but an activity involving someone who is doing the thinking/feeling/perceiving...

These comments are, I recognise, all fairly negative, tending towards the restoration of ordinary, unexciting common sense. But that was the tendency of philosophy in the 50s - which may help to explain why so much high-powered effort seems to have produced so little of lasting philosophical importance. 'Getting the fly out of the fly-bottle' is a relatively modest ambition.