TONY CHADWICK ESSAY PRIZE 1999 WINNER

THE MIND IS NOT THE BRAIN: THE IRREDUCIBILITY OF THE MENTAL

by Phil Rees

1. INTRODUCTION

In the 1950s various philosophers proposed a form of mind-brain identity[1]. Their proposals came to be known as the Mind-Body Identity Theory. The theory held that mental states are contingently identical with states of the central nervous system (which of course includes the brain). The word 'contingently' here is important since the identity theory was proposed as an empirical discovery, not as a logical necessity arising from the meaning of terms. I shall argue that the identity theory is false.

2. THE IMPOSSIBILITY OF VERIFICATION

The identity between mental and brain events is stated to be contingent and therefore requiring empirical verification. However, in principle, no experimental method could ever be devised to verify the proposition of the identity theory. The theory states that brain state S_b is in fact identical with mental state S_m. Suppose some kind of probe is inserted into the brain to establish that brain state S_b occurred at some particular time. How will it be established that mental state S_m also occurred at that time? No means exists for the detection of mental events or states other than by being reported by the subject concerned. So we would have to ask the subject to verify that he was in mental state S_m at that particular time. But if he attempted to do this then he would not be in mental state S_m, but in some other state arising from the fact that he was trying to give a confirmation (by whatever means this was to be done). His mental state would be perturbed and we would not find what the experiment might have hoped for. Examination of mental states has indeterminacy built-in. Any attempt to probe the subject changes the conditions of the experiment. So the contingent theory that mental states are identical with brain states is in principle unverifiable. We do not need to be logical positivists to feel that any scientific proposition which is unverifiable in principle is not in fact a scientific proposition at all, but a statement of faith.

3. LACK OF EXPLANATORY POWER

Let us suppose that the identity theory is in fact correct and that there is an identity between mental states and brain states. We now wish to use this fact to add to our ability to explain human behaviour. We ask 'Why did you pick up that pen?' Do we expect an answer in terms of underlying brain states? Could any such description in the language of brain states ever lead to the answer we are really interested in: 'Because I wished to write a letter to my friend'. The explanatory language of mental states is at the macro level and no reference to the micro level will have any explanatory power. An analogy can be drawn with a large tank of water in which scientists are studying hydrodynamic wave flows. It is true to say that the water consists of H_2O molecules - the 'water is H_2O'. Every possible configuration of the waves in the tank (macro states) at any given time is identical with some state of the molecules (micro states). Nevertheless, no examination of the micro states could ever give the kind of macro level explanation we would be looking for. The fact that the macro structure is the micro structure doesn't give any explanatory power at the macro level. The dynamics of wave motion in the tank is explained in terms of the macro properties (such as the dimensions of the tank), not the micro properties.

In the same way, the dynamics of the macro structure that is the mind is explained in terms of the macro properties, not the micro properties - the brain states which (let us suppose) underlie it. So we have an empirical theory that can never be verified and which, even if it were true would carry no useful explanatory power.

4. IDENTITY OR REDUCTION?

The proposition under consideration is 'Mental states are identical with states of the brain'. Notwithstanding the above negative points, it seems to me that the key issue is an analysis of what is meant by identity. What does the phrase 'are identical with' mean?

Since the identity is claimed to be contingent we are not dealing with the kind of identity displayed by the statement 'all bachelors are single' since this states a necessary and analytic identity. Smart seems unclear as to exactly what kind of identity is being considered[2]. Despite stressing strict identity, he gives as an example the statement '7 is identical with the smallest prime number greater than 5' which is certainly not contingent.

Perhaps the kind of identity intended is that displayed by the statement 'the morning star is the evening star'. This is certainly the kind of strict referential identity that Smart tells us he intends. However, if this is the case then the identity theory will be subject to Leibniz' law.

5. IDENTITY AND LEIBNIZ' LAW

I propose any claim of strict identity between mental states and brain states cannot be compatible with Leibniz' law of the indiscernibility of identicals. This states that if two things, x and y, are identical, then every property of x must also be a property of y and vice versa (excluding intentional and modal properties). There are several properties of mental events and brain events which violate this law. It should be stressed that all of these examples are perfectly in accord with a reductive identity, but not with a strict referential identity.

5.1 Physical location

It is hard to see how the physical location of brain events and mental events can be the same, or even how physical location of mental events can be coherently considered. We cannot assign any kind of spatial location to thoughts. What of sensations (aches, twinges, etc.)? These don't really have definite spatial locations. However, it might be argued that they are specifically located. We might feel that a pain in the left toe is located in the left toe. If this is so then it is certainly not located in the brain.

If by some means we managed to establish that at the moment I had a particular mental event E_m, brain event E_b was occurring at a particular place in my brain, what independent tests could we even conceive of that might verify that mental event E_m was occurring at the same place in my brain?

5.2 Epistemic asymmetry

As discussed above, if we wish to determine whether a brain event has occurred we can do so using scientific instruments. But how do we make any inspection of mental events? By consulting the subject, which in turn alters the results of the inspection. In terms of the properties of inspection, measurement, etc., the two sides of the supposed identity could hardly be more different.

However, it goes deeper than this: The plain fact is that knowledge of mental events cannot be objective. There is no external epistemic access to mental events as there is for brain events. The identity theory asserts identity between events that are epistemically objective and events that are epistemically subjective.

5.3 Explanatory language

As discussed above, the language and explanatory system of brain states is completely different from that of mental states. A typical statement about brain states would be something like 'The a region neurones showed 5Hz/8Hz/11Hz spikes which led to the arousal of C-fibres in the ß region'. A typical statement about mental states might be 'My jealousy caused me to be angry, so I struck him'. This kind of acute difference in the language of two sides of an identity is the hallmark of a reductive identity, but is not compatible with the claim of strict identity.

6. REDUCTIVE IDENTITY

My thesis here will be that, if the relationship between mental and brain events is that of reductive identity, then it must be a causal reduction and not an ontological reduction.

There is much evidence to suppose that, although Smart tells us he intends 'strict identity'[3], the theory is actually founded on some kind of reductive identity. Most of the papers on the identity theory cite scientific reductive identities as paradigm examples.

Let us consider reductive identity. This is discussed with exemplary clarity by John Searle[4.] His analysis draws out five different types of reduction, but, for our purposes, the important two are ontological reduction and causal reduction. Here are some examples of contingent identity statements taken from the essays of Place, Smart, and Feigl, and cited by them as examplars of the kind of identity they have in mind in the identity theory:

(a) Water is H_2O (ontological reduction).
(b) Heat is mean kinetic energy of molecules (causal reduction).
(c) Lightning is electrical discharge (causal reduction).

My assertion is that none of these are concerned with strict identity but are examples of reductions.

To take the first case, water cannot be identified literally with H_2O. H_2O is a formula for describing the way in which a particular kind of molecule is constituted - namely of two atoms of hydrogen and one of oxygen. So the best we could make of (a) is 'the chemical composition of water is H_2O molecules' and this is certainly not an identity statement. In this case it is not the 'is' of identity but the 'is' of composition. Clearly this is an ontological reduction from 'water things' to 'molecule things'.

Heat is (or was) an essentially phenomenal term. It is a word coined by humans to describe their perception of a natural phenomenon. Mean kinetic energy of molecules is a number describing the extent to which the molecules in a substance are moving. Strict identity is clearly not involved. The best we can say is 'the phenomenon of heat is caused by the mean kinetic energy of molecules'. This is a causal reduction.

Lightning, again, is an essentially phenomenal term, and describes human perception of a natural phenomenon. Strict identity is again not involved, and we have another causal reduction.

Searle points out that most cases of scientific causal reduction are accompanied by subsequent ontological reductions[5]. In the case of water and H_2O it may be the case that, over time, the 'is' of composition gets transformed into the 'is' of identity by a subtle transformation of the usage of words, so that in most contexts 'H_2O' ceases to be a chemical formula but becomes an alternative way of referring to the stuff that 'water' refers to. When this happens 'water is H_2O' has become a strict identity[6].

Prior to the reduction of heat to 'the kinetic energy of molecules' the word 'heat' included the subjective feeling - the 'raw feel'. Having constructed a causal reduction that explained the cause of that subjective 'heat', in some contexts the word effectively changed its meaning so as to refer to the underlying objective phenomenon. The causal reduction led to an ontological reduction.

It is interesting to note that, with each of these reduction examples, when we move from the terminology of the reduced entity (e.g. heat) to that of the reducing entity (mean kinetic energy), whilst we gain in terms of our analytical understanding of the physical constituents, we lose in terms of subjective, phenomenal descriptive power. The phenomenal nature of natural phenomena cannot be reduced. The 'raw feels' do not reduce.

7. THE IRREDUCIBILITY OF CONSCIOUSNESS

We reach the central problem for the identity theory - the ontological irreducibility of conscious phenomenal experience. There are many ways in which this irreducibility has been demonstrated by different writers.

7.1 Nagel on subjective/objective

Nagel says that '... an organism has conscious states if and only if there is something it is like to be that organism - something it is like for the organism'[7] He accepts that the phenomenon of consciousness may be caused by physical brain states, but denies that any such analysis of consciousness can be complete. Any reductive analysis omits the subjective character of conscious experience. All attempts so far at reductive analysis have been equally compatible with the absence of conscious experience.

He dwells on the asymmetry between objective and subjective, making the point that, whilst it is correct to objectively examine a phenomenon like lightning in an attempt to reduce it to electrical discharge, the idea of objectively examining experience is not coherent. Experience can be examined only by the organism having it. '... what would be left of what it is like to be a bat if one removed the viewpoint of the bat?'[8].

7.2 Jackson and Robinson on phenomenal knowledge

Jackson's 'knowledge argument'[9] is effective against any attempt to reduce the mental to the physical. The now familiar argument concerns Mary, a neuroscientist who knows all the physical facts about the brain's perception of colour. However, Mary has spent her entire life in a black and white environment. When she is released and sees a red rose she learns what it is like to see red, and this is new phenomenal knowledge for her. The argument concludes that the

physical facts do not exhaust all the facts, and consequently the mental cannot be reduced to the physical.

Robinson[10] makes a similar and equally convincing argument concerning a deaf neurologist who possesses all the facts about hearing but gains new phenomenal knowledge when his deafness is cured.

7.3 Lockwood on what it is like to be another person[11]

Lockwood points out that whilst the above arguments are convincing, they have been criticised on the grounds that those concerned only lack *concepts*[12]. Mary does in fact have all the *knowledge* she needed; she just lacked the *concept* of red. Similarly, in Nagel's argument we lack the *concept* of the phenomenology of a bat's sonar. Lockwood grants that it could be argued that

> '... it is unreasonable to regard as an obstacle to reductionism, the impossibility of deriving certain conclusions from the physical descriptions when one lacks the relevant higher-level concepts. For that would have the effect of rendering it impossible to ever carry through any scientific reduction.'[13].

He goes on to make what seems to be the most conclusive argument against the reduction of the phenomenal: He poses the question - suppose we had total physical knowledge of everything about another person and that we have unlimited intellectual powers; could we now know what it is like to be that person? Clearly, the 'concept' argument cannot now be used because we have all the right concepts. Could we know what it is like for that person to feel a particular pain which he claims he feels? No, we could not because we would not know how to interpret all these physical facts. If we had a similar total physical description of ourselves, we could seek a correlation with what it is like to be ourselves in various similar physical states and hence draw conclusions about what it is probably like to be the other person. However, to do this would be to draw on our own subjective and phenomenal knowledge, information that goes beyond that contained merely in the physical facts. Consequently, the attempted reduction to the physical facts loses phenomenal knowledge.

7.4 The difference between heat and consciousness

Why should we find that heat can be successfully reduced to mean kinetic energy of molecules, but consciousness (or phenomenal experience) cannot. Referring to our discussion of heat above, we saw that first there was a causal reduction and this was followed by an ontological reduction.

We can also make the causal reduction for consciousness. We can regard consciousness as a macro property of the brain caused by the micro properties of neurones, C-fibres, etc. This is a causal reduction. However, to follow this with an ontological reduction involves redefining 'consciousness' to refer to the micro properties and we then simply lose the subjective content which consciousness is all about. As we observed above, the ontological reduction that follows any causal reduction always tends to lose the phenomenal aspect of what is being reduced. But to lose the phenomenal aspect of consciousness is to lose consciousness itself.

The contrast between the reducibility of heat and the irreducibility of consciousness is only apparent. When we applied an ontological reduction to heat, we did not really lose the

subjectivity of heat, but simply stopped calling the subjective part 'heat'. We did not eliminate any subjective phenomena, but just stopped calling them by their old names.

8. CONCLUSION

I have criticised the Identity theory on the grounds of the impossibility of its verification. I then also examined the nature of its claims of identity between mental states and brain states. It has been shown that if the intention is strict identity then Leibniz' law is contravened in several fundamental ways. On the other hand, if a reductive identity is intended then it is argued that it can only be a causal reduction since any attempt at ontological reduction leaves conscious, phenomenal experience out of the reduction.

Perhaps we should recall the words of Bishop Joseph Butler, 'Everything is what it is and not another thing'[14]. The mind is not the brain.

References:

1. Place, 1956; Feigl, 1958; Smart, 1959
2. Smart, 1959, p.56.
3. Smart, 1959, p.56
4. Searle, 1994, pp. 112-6
5. Searle, 1994. p. 115
6. Presumably this had happened by the time Kripke's *Naming and Necessity* was written since he in fact uses 'water is H_2O' as an example of a strict identity between rigid designators I stand by my argument nonetheless.
7. Nagel, 1974, p.160
8. Nagel, 1974, pp.163-4
9. Jackson, 1982 and defended in Jackson, 1986
10. Robinson, 1982
11. Lockwood, 1989
12. Churchland, 1985
13. Lockwood, 1989, p.132
14. Butler, 1736

Bibliography

Butler, J, 1736, *On Personal Identity* first published as an appendix to *his The Analogy of Religion Natural and Revealed to the Constitution and Course of Nature*

Churchland, P.M., 1985, 'Reduction, Qualia, and the direct introspection of brain states' in *Journal of Philosophy*, 82. Reprinted as chapter 3 of his *A Neurocomputational Perspective* (MIT,1989)

Feigl, H., 1958, 'The "Mental" and the "Physical"' in (Feigl, Scriven, Maxwell, eds.) *Minnesota Studies in the Philosophy of Science* vol. 2*: Concepts Theories and the Mind-Body Problem* (Univ. of Minnesota Press, 1958). Also reprinted in many other anthologies.

Jackson, F., 1982, 'Epiphenomenal qualia', *Philosophical Quarterly* vol. 32. Reprinted in (W.Lycan, ed) *Mind and Cognition*, (Blackwell, 1990)

Jackson, F., 1986, 'What Mary Didn't Know', *Journal of Philosophy* 83:5. Reprinted in (Block, Flanagan, Guzeldere, ed.) *The Nature of Consciousness*. (MIT, 1997)

Lockwood, M., 1989, *Mind. Brain and the Quantum*. (Oxford, 1989)

Nagel, T., 1974, 'What is it like to be a bat?' in *Philosophical Review* 83. Reprinted (Block, ed.) *Readings in Philosophy of Psychology*. (Harvard, 1980). Page numbers quoted are from Block.

Place, U.T., 1956, 'Is consciousness a brain process?' *British Journal of Psychology* vol. 47. Reprinted in (W. Lycan, ed.) *Mind and Cognition* (Blackwell, 1990)

Robinson, H., 1982, *Matter and Sense*.(Cambridge, 1982)

Robinson, H., 1993, 'The anti-materialist strategy' in (H. Robinson, ed.*) Objections to Physicalism.* (Oxford, 1993)

Ryle, G., 1949, *The Concept of Mind*, Penguin

Searle, J., 1994, *The Rediscovery of the Mind*, (MIT press, 1994)

Smart, J.J.C., 1959, 'Sensations and Brain Processes' in *Philosophical Review* vol. 68. Reprinted in (Borst, C.V. ed.) *The Mind-Brain Identity Theory* (Macmillan, 1970)