

CONTROVERSIES IN PHILOSOPHY

General Editor A. G. N. Flew

THE IS-UGHT QUESTION

Edited by W. D. Hudson

THE MIND-BRAIN IDENTITY THEORY

Edited by C. V. Borst

PHILOSOPHY AND LINGUISTICS

Edited by Colin Lyas

WEAKNESS OF WILL

Edited by G. W. Mortimore

THE PRIVATE LANGUAGE ARGUMENT

Edited by O. R. Jones

The Mind-Brain Identity Theory

A COLLECTION OF PAPERS
COMPILED, EDITED AND FURNISHED
WITH AN INTRODUCTION BY

C. V. Borst

CONTRIBUTORS

*D. M. Armstrong, Kurt Baier, Robert Coburn, James Cornman
Herbert Feigl, Paul Feyerabend, J. M. Hinton, Norman Malcolm
Thomas Nagel, U. T. Place, Richard Rorty, Jerome Shaffer
J. J. C. Smart, Ernest Sosa, J. T. Stevenson, Charles Taylor*

MACMILLAN

© Selection and editorial matter C. V. Borst 1970

All rights reserved. No part of this publication
may be reproduced or transmitted, in any form
or by any means, without permission

First edition 1970
Reprinted 1973 (twice)

Published by
THE MACMILLAN PRESS LTD
London and Basingstoke
Associated companies in New York Dublin
Melbourne Johannesburg and Madras

SBN 333 10039 5 (hard cover)
333 11047 1 (paper cover)

Printed in Great Britain by
WESTERN PRINTING SERVICES LTD
Bristol

The paperback edition of this book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser

Foreword by the General Editor

The series of which the present volume is the second member is to consist largely but not entirely of material already published elsewhere in scattered sources. It is as a series distinguished by two guiding ideas. First, the individual editors of the various constituent volumes select and collect contributions to some important controversy which in recent years has been, and which still remains, alive. The emphasis is thus upon controversy, and upon the presentation of philosophers in controversial action. Second, the individual editors are encouraged to edit extensively and strongly. The idea is that they should act as firm, fair, and constructive chairmen. Such a chairman gives shape to a discussion and ensures that the several contributors are not merely heard, but heard at the moment when their contributions can be most relevant and most effective. With this in mind the contributions as they appear in these volumes are arranged neither in the chronological order of their first publication nor in any other and arbitrary sequence, but in such a way as to provide and to reveal some structure and development in the whole argument. Again, and for similar reasons, the editorial introductions are both substantial and forthcoming.

They can be seen as representing a deliberate rejection, at least within this special limited context, of the 'throw-a-reading-list-at-them, send-them-away, and-see-next-week-whatever-they-have-made-of-it' tutorial traditions of some ancient British universities.

The problem to which the Mind-Brain Identity Theory is offered as a solution was set by Descartes. For it was Descartes who persuaded modern philosophy to put enormous weight upon a fundamental distinction between mind and matter, consciousness and stuff. The problem is to say how the two can be, and are, related. Among the traditional answers have been epiphenomenalism, parallelism, and interactionism. The first of these urges that consciousness is some sort of result or property of a certain sort of material thing, but is incapable of any reciprocal effect on matter. Among the analogies

Editor's Introduction

C. V. Borst

I. CONTINGENT IDENTITY

The papers in this volume are all concerned with expounding and examining a certain view of the nature of the relation between mind and body, mental states and physiological states. The particular theory under discussion is commonly known as the Identity Theory, though at least some versions of it also go under the name of Central-state Materialism.

The view is, as the name suggests, that mental states are quite literally identical with brain states: any given mental state is, roughly, a brain state, brain process or feature of a process in the central nervous system. However, what is distinctive about the currently canvassed Identity Theory is that the proposed identity is put forward as a scientific discovery – or at least potential scientific discovery – and not as a truth concerning the meaning of mental terms or concepts. This explains a good deal.

At first sight it might seem strange that so apparently obvious a suggestion as this identification has not, until so recently, been much discussed by philosophers. The reason, though, is not hard to find. The suggestion has been thought open to insuperable objections. Philosophers, especially since Descartes, have usually defined the mental in terms incompatible with the physical; mind for Descartes was in essence unextended, matter in essence extended. Even at a more everyday level the mental is customarily contrasted with the physical – rather as hot is exclusively contrasted with cold, wet with dry, light with dark, abstract with concrete. Any suggestion, therefore, that the mental is physical inevitably carries with it an air of paradox, if not of outright contradiction; for it seems, on the face of it, to imply that the mental is not mental.

However it is certainly now common knowledge that the brain, even although a physical organ, bears some intimate relation to the mind; and this knowledge is manifestly embodied in common idiom. When referring to admittedly mental happenings or conditions we speak of racking one's brains, picking another's brains, having it on the brain, having a good brain or, contrariwise, being all brawn and

no brain; and so on. Hence the question inevitably arises of just what the relation is between mind and brain. Few would want to dispute that the brain has a great deal 'to do' with the mind. Injuries to the former are known to impair functions associated with the latter; certain drugs are presumed to alter mental states by their effect on the electrochemical processes of the nervous system; psychologists and physiologists have discussed the localisation of mental functions in specific regions of the brain. Even Descartes was prepared to allow, with doubtful consistency, a special role, in relation to mental operations, to the brain and especially to the pineal gland.

An obvious enough proposal is that the relation is of a causal nature. Sensations are, perhaps, the effects of stimulation of the sensory areas of the brain; volition perhaps causes motion of the limbs by its effect on the motor areas. Since it only makes sense to speak of causal transactions between ontologically distinct phenomena, the result is a dualist point of view. The difficulties of attempting to explain the causal interaction of fundamentally unlike phenomena have, notoriously, led to various psycho-physical parallelist doctrines. An obtrusive feature of any proposal to *identify* the mental with the cerebral is, manifestly, a denial of any form of psycho-physical or more specifically of mind-brain dualism, and an assertion of some form of philosophical monism. If the mental is still, none the less, conceived as a special category it will be so only as a particular subclass of the physical. Indeed just such a point of view is characteristic of what can profitably be seen as the predecessor of the Identity Theory, an alternative monistic account, namely, Behaviourism.

II. BEHAVIOURISM

Significantly the two philosophers chiefly responsible for the introduction of the Identity Theory, Professor Herbert Feigl and U. T. Place, both take as starting points of their discussions some form of behaviourist doctrine. Behaviourism, in fact, started out as a psychological theory. It was originated in the second decade of the present century by J. B. Watson. Instead of the predominantly introspective study of states of consciousness, psychology was to become the scientific study of human and animal behaviour. What then becomes of states of consciousness? These can occur in the absence of observable behaviour, so it looks as if the two cannot simply be equated, at least on the face of it. One way of meeting this point was by means of the notion of *covert* behaviour. Thinking, for example, was equated with or replaced by minute movements

of the larynx or laryngeal muscles, and perhaps also of other organs associated with speech. Yet it could hardly be maintained that what was commonly *meant* by 'thinking' consisted in, or indeed had anything at all to do with, laryngeal movements. Rather the suggestion was that what went under the name of 'thinking' was *as a matter of fact* nothing but certain minute movements of the larynx (or in the case of deaf mutes, of the muscles of the fingers). Less ambitiously, but more plausibly, it was alternatively maintained that for the purposes of psychological investigation thinking could best be *regarded as* larynx movement: thus conceived, Behaviourism was a methodological principle. Psychology should, in common with the established sciences, deal exclusively with intersubjectively observable phenomena and not rely on non-intersubjectively confirmable reports of occurrences 'in the minds' of its subjects.

In contrast to this methodological Behaviourism and to the bolder doctrine that might be called *Contingent* Behaviourism (but which is otherwise known as *Metaphysical* Behaviourism), was a doctrine developed by philosophers which is called *Logical* (or *Analytic*) Behaviourism. Here the viewpoint is one not of science but of logical or conceptual analysis. The contention was that the meaning of mental statements was analysable, without remainder, into statements about behaviour (possibly also about physiological changes) and about the observable circumstances in which such behaviour occurred. The previously mentioned objection that mental processes can occur in the absence of any overt behaviour was taken care of, not by invoking the notion of covert behaviour, so much as by the notion of unfulfilled disposition to behave. Thus someone may be angry whilst sitting apparently undisturbed in an armchair yet, it would be said, he is none the less necessarily disposed to act aggressively, and would do so but for willed or learned restraint. Such a doctrine, with very great refinements, is commonly associated with the names of Rudolf Carnap, Gilbert Ryle, and, some would contentiously add, Ludwig Wittgenstein. A close parallel as regards the type of thesis concerned is provided by Logical (or Analytic) Phenomenalism, to the effect that any statement about a physical object can be analysed into a set of statements about sense-impressions. In both cases the logical tool involved can be seen as Bertrand Russell's notion of a logical construction. The idea here is, roughly, that A's are logical constructions out of B's if and only if all statements about A's can be replaced, without loss of meaning, by concatenations of statements about B's. Physical objects become logical constructions out of sense-impressions, minds out of behaviour,

somewhat as nations might be logically constructed out of individual citizens or universities out of their several colleges or departments. But in no case was any real entity to be postulated over and above the constituent entities. Russell's own principle was that where possible, logical constructions should replace inferred entities, on the ground that avoiding the unnecessary postulation of unobservable entities reduces the risk of being mistaken, provided that one does not explicitly deny that such entities may exist. However, this principle does tend to overlook the explanatory advantages of judicious entity-postulation and the fact that through technological advances one-time unobservable entities can become, at a later time, observed, in some sense or other.

III. CAUSAL EXPLANATIONS AND THE IDENTITY THEORY

U. T. Place begins his discussion by admitting that a dispositional analysis is fundamentally sound in the cases of our cognitive and volitional concepts, but he contends that there remains a number of our mental concepts, of consciousness, of sensation, and mental imagery where such an analysis fails, and an inner process account becomes unavoidable. Place was followed in this by Professor J. J. C. Smart, who originally applied the Identity Theory only to the case of sensations. Smart asserts that his 'Sensations and brain processes' takes its departure from Place's 'Is consciousness a brain process?' while also supplementing Feigl's 'The "Mental" and the "Physical"'. But it is in fact Smart's own article which seems more than any other to have caught the attention of philosophers and to be primarily responsible for putting the Identity Theory firmly on the current philosophical map. Since its publication in 1959 this relatively brief article has been veritably echoing through the pages of the journals in the form of discussion which it has engendered, and must, in addition, surely be near to creating a record as the most frequently reprinted article.

Feigl, also, in his mammoth essay 'The "Mental" and the "Physical"' (see Bibliography) develops his theory from a consideration of Behaviourism, though in this case the primary concern is with the molar behaviour theory of the psychologist (a résumé of his position is given in the included article, Paper I). Feigl points out that theoretical concepts are frequently employed by psychologists in the explanation of behaviour. Certainly there are some psychologists, of whom B. F. Skinner is a notable example, who wish to work solely in terms of observables. However Feigl's argument is that as soon

as one passes, as he clearly believes should be done, beyond this *peripheralistic* Behaviourism and introduces *central* states in the explanation of overt behaviour, then the way is opened to a twofold identification. Firstly these central causes are identified with the referents of the phenomenal terms used by subjects in introspective descriptions of direct experience, and secondly these postulated central states are further identified with the referents of certain neurophysiological terms.

These postulated central states are in fact an example of the psychologists' so-called *hypothetical constructs* and as such are not definable in terms of observables, though they are logically related to such observables. Being thus ontologically distinct states they are able, as previously indicated, to feature in causal explanations of observable behaviour. The contrast is with the misleadingly called *intervening variables*, which are similar to the philosophers' logical constructions. Here there is no postulation since the meaning is restricted to observables, as the meaning of the term 'thirst' would be if it were taken to mean no more than, say, not having had a drink for a considerable time, and there thus subsisting a high probability of drinking-behaviour if presented with the opportunity.

Now one of the merits which proponents of the Identity Theory are wont to claim as against Logical Behaviourism is that their theory enables them to assign to mental states a genuinely causal explanatory role. This, they contend, accords with the role that such states carry in ordinary thought and speech. Take again the case of anger. As conceived by the Identity Theory, anger is thought of as an inner state over and above and standing behind angry-behaviour and bringing it about, so that it then becomes possible to construe such everyday remarks as 'He shouted *because* he was angry' or 'His anger made him go red in the face' in a genuinely causal sense. Whereas if anger is conceived as a logical construction out of angry-behaviour the explanatory force of such remarks is confined to being of a generalisatory nature. To put it briefly, he shouted or turned red because he was in a state constituted by being disposed, or possessing a tendency, to shout, turn red, and generally act aggressively: and not because he had difficulty in making himself heard, was blushing, or whatever.

One interesting corollary of the causal account is that it seems to provide a possible way of understanding, at least in the case of some psychological concepts, the puzzling relation between a criterion, in Wittgenstein's sense, and what it is a criterion for. The difficulty

encountered in attempting to characterise this criterial relation has always been to specify a logical relation which is other than that of straight logical equivalence. In fact Armstrong draws attention in his book (*A Materialist Theory of the Mind* (1968) p. 92) to the striking fact that Wittgenstein's dictum, 'An "inner process" stands in need of outward criteria' could serve as the slogan of a causal analysis of mental concepts. Although on his theory the quotation marks could be removed from the words 'inner process'. Consider the most general characterisation of a mental state which Armstrong advances: it is, he says, 'a state of the person apt for producing certain ranges of behaviour' (p. 75, below). This characteristic behaviour would then become the criterion for the ascription of the relevant mental state.

Of course it will be tempting for a critic of the Identity Theory to complain that its suggested explanatory concepts are really little better than the notorious 'explanation' of the soporific power of opium in terms of a certain *virtus dormitiva* – it is due to a dormative principle. There does indeed appear to be something in this criticism. But it can in part be met if the alleged explanatory concept forms part of a theoretical framework whose overall explanatory power is capable of explaining and predicting further phenomena. Furthermore the position is strengthened by making reference to the second stage of the two-stage exposition of the Identity Theory as it is given by Smart and Armstrong. The first stage attempts to exhibit our mental concepts in such a way as to show their *compatibility* with a materialist metaphysic: this is seen as a task for purely philosophical analysis. The general form of Armstrong's stage-one account has already been indicated. The earlier account of Smart, which differed from that of Armstrong in stressing the causes rather than effects of mental states, was actually exclusively concerned with an analysis of sensation-reports, in what he described – in Ryle's phrase – as 'topic-neutral' language. To report, for instance, the occurrence of a yellowy-orange after-image amounted, to a first approximation, to reporting that 'something is going on in me which is like what goes on when I have my eyes open and there is an orange in front of me'.

Such schemata are not held actually to entail a materialist account of mental states in that the 'state of the person apt for . . .' or what 'is going on in me when . . .' could in fact turn out to be psychical, spiritual, mentalistic, in short, a non-physical occurrence. But then the claim of the second stage of the argument is that, as a matter of contingent fact, and on the basis of present scientific knowledge, it

is most plausible to identify such occurrences with states of the central nervous system. This second stage is thus one not of philosophical analysis but of empirical fact.

A favoured model is the case of the scientific identification of the gene with the D.N.A. molecule. A gene is characterised in such a way as to be logically linked with its observable effects, the transmission of hereditary information. Exactly what the intrinsic nature was of the entities responsible for this transmission was at the time of their initial postulation unknown. But now, much later, the evidence for their identification with the D.N.A. molecule has become overwhelming. (Looked at in this light the assignment of a certain *virtus dormitiva* to opium might seem considerably less nugatory than would otherwise appear – it could be seen as the positing of some sort of structure or ingredient of opium, of a then unknown nature, possessing the observed soporific effects.) The general line of thought is that many of our ordinary explanatory concepts contain a sort of gap or lacuna which later empirical findings are able to plug.

IV. ALTERNATIVE FORMS OF THE IDENTITY THEORY

An indication can now be given of the various different forms of the Identity Theory which have been put forward and which are represented in this collection. On the one hand there are the various antipodean versions from what Feigl has called the 'United Front of Sophisticated Australian Materialists', notably Smart and Armstrong, though the precursor of this position was the exposition of U. T. Place (written while he was a member of Smart's department at Adelaide). Their outstanding characteristic is a tough-minded and uncompromising materialism: pains, thoughts and after-images exist – or at least the experience of having them does – but are in fact nothing but states of the central nervous system of the person who has them. As originally put forward by Place and Smart, the central-state materialist account was to apply only to that limited set of mental concepts considered resistant to a behaviourist treatment. But with the appearance of Armstrong's exposition it was extended to cover every one of our mental concepts, and Smart now goes along with Armstrong in this. This extension has, however, generated additional powerful objections.

Then there is the quite considerably different form of the theory put forward by Feigl. This is harder to classify. But it can still be seen as a different version of basically the same theory in that it,

too, concerns a *contingent*, as opposed to a *logically necessary*, identification. Feigl was perhaps the first exponent of the theory to make explicit use of Frege's distinction between the sense of an expression and its reference, and of the utilisation of Frege's model of the empirically discovered identity of the Morning Star and Evening Star. In taking the common referents of phenomenal terms and certain neurophysiological terms (having, of course, logically independent senses) to be *raw feels*, Feigl might seem to be advancing a purely idealist version of the Identity Theory, as opposed to the Australian materialist version. But in point of fact Feigl sees himself as defending a *physicalist* position. (The term 'physicalist' has customarily been used for the view that every statement can be expressed in the language of physics.) This is so for Feigl because of his belief that 'the basic *laws* of the universe are the *physical* ones' (p. 40, below): all occurrences are related by and explainable in terms of the laws of the physical sciences (with the possible need to include 'emergent' laws in the case of biological phenomena). The term 'raw feels' was coined by the psychologist E. C. Tolman: by it is meant, roughly, the unconceptualised items of direct experience, sentience, the phenomenally given (e.g. sense-data or sensations). Feigl also appears to be following a suggestion of Tolman in construing such raw feels as the 'realities-in-themselves' which the neurophysiological terms denote.

Now there is, in addition to the above, a third form of the Identity Theory which has been dubbed by Richard Rorty the 'disappearance form', in contrast to the more familiar 'translation form', so called in virtue of the stage-one 'translation' of our ordinary mental concepts into topic-neutral language. It is this disappearance version which Rorty himself defends in the included article (Paper XX); and Professor P. K. Feyerabend's theory (see Paper XV) can also be classified as a 'disappearance' theory. The distinctive feature of this form is that no attempt is made to defend our ordinary thought and speech. On the contrary, the incompatibility of our ordinary mental concepts with scientific discoveries, and projected discoveries, is stressed: sensations and sensation-talk, for example, will simply disappear from a scientifically oriented language, much as demons and demon-talk have already disappeared. Their place might be taken by 'successor' concepts, which, nevertheless, would not stand in any simple relation to the present concepts. Not that it is necessarily suggested that this replacement would actually happen in practice, but in so far as it did not this would be largely because of evident inconvenience. We would, I suppose, for a while at least, be permitted, as

Berkeley allowed, to 'speak with the vulgar' so long as we were to 'think with the learned'.

Doubt has, as a matter of fact, been raised (see James W. Cornman, 'On the Elimination of "Sensations" and Sensations', *Review of Metaphysics*, xxii (1968) pp. 15-35) whether the disappearance-theory is correctly termed an *identity* theory at all. But undoubtedly the general import is very similar and, moreover, as Rorty claims, the *is* of 'strict' identity can still perfectly properly be employed in statements of the form: 'What people now call "sensations" are (identical with) certain brain processes.' Significantly, in the face of various objections to his initial position, Smart confesses himself attracted to the disappearance or replacement theory of Feyerabend (see the final paragraph of 'Materialism', below) and Feigl, too, in his 1967 Postscript to 'The "Mental" and the "Physical"' shows some tendency to move in a similar direction.

V. IDENTITY STATEMENTS

The truth is that there has from the start been much difficulty in providing a coherent account of the proposed identity between mental states and neurophysiological states. Largely this arises because of the logical symmetry of *bona fide* identity statements: if A is identical with B, then equally B must be identical with A. Moreover if the identity is to be logically contingent, independent 'definitions' of the two terms must be provided. In Frege's example of the Morning and Evening Stars, the two terms on either side of the identity-sign possess equal status; there is no question of the Morning Star being *really* the Evening Star in any sense in which the Evening Star is not *really* the Morning Star. Worse still, on the translation Identity Theory mental states were to be *really* brain processes without brain processes being *really* mental states, but for all that the mental states were to remain unscathed! The following illustration is possibly instructive. Suppose for the sake of argument that the proposition 'Shakespeare is Bacon' is true; then someone might be inclined to remark that, in that case, there is no such person as Shakespeare (as commonly conceived) and so how can *he* be Bacon? Just such a point tends to arise for the Identity Theory; if mental states are, despite appearances, brain processes then there might seem to be no mental states to *be* brain processes. If the theory is true then it is false. To which a disappearance theorist can reply that that is precisely so; there are no mental states as commonly conceived. What people have misguidedly called 'mental states'

are really brain processes, and what people have been calling 'Shakespeare' is in fact Bacon (in one of his roles, or wearing the particular 'hat' of the writer of the 'Shakespeare' plays). All the same this move needs to be treated with the greatest circumspection, for in some cases the results of applying it can be palpably absurd, as, for example, in the case of knowledge. Any argument along the lines 'what people now call "knowledge" is really x, y, z, therefore we should cease talking of knowledge, for there is something deficient about the concept' seems to me to possess no merit whatever. This may be because the term 'knowledge' carries little if any theoretical load and is, in any event, certainly not a straight referring expression.

Now the translation form of the Identity Theory would not suggest that the concept of knowledge might disappear, for there the suggestion was not that our ordinary psychological concepts are defective but that they are somehow unspecific. The topic-neutral analysis was in fact put forward in order to avoid the need to give an independent identity to one side of the alleged identity statements. In seeking models, proponents have characteristically turned to instances of contingent identities apparently found in the natural sciences. Such statements as 'Water is H_2O ', 'Lightning is an electric discharge', 'Heat is mean kinetic energy', 'Solubility is the possession of such and such a structure', 'Genes are D.N.A. molecules' have been suggested. These certainly appear promising; for are they not genuine identity statements which, nevertheless, have an asymmetrical nature, being reductive in the sense that the right-hand side specifies the discovered nature of the left-hand side? The term 'theoretical identity' has sometimes been employed, for example by Feigl, and by Hilary Putnam (in 'Minds and Machines', *Dimensions of Mind*, ed. Sidney Hook (New York 1960) pp. 148-79). However the position is far from straightforward, because it can well be doubted both whether the scientific cases are indeed genuine identity statements, and whether they do really furnish viable models for the required theory of mind. The examples purport to give the physico-chemical or micro-structure of the things or properties in question, and it is possible to read them as providing *explanations* of familiar phenomena rather than as specifying what the things or properties *really are*. Thus it could be argued that reference to the mean kinetic energy of component molecules supplies an explanation of the familiar phenomena of heat and that heat itself should still be defined in terms of the boiling of kettles, the melting of wax, the burning of the skin, and so forth; or that the property of solubility should still be defined in terms of the fact that a substance which is

soluble will dissolve when placed in water, this being explained by, rather than being, a certain physico-chemical structure.

It is doubtful if there is exactly any right or wrong thing to say in this dispute. But an interesting point is made by Putnam in a later article ('The Mental Life of Some Machines', *Intentionality, Minds, and Perception*, ed. Hector-Neri Castañeda (Detroit 1967) pp. 177-213). Putnam considers that it would only not be a 'wholly unmotivated extension of ordinary usage' to say that solubility is the possession of a particular physico-chemical composition if *that* particular composition is physically necessary to the property of solubility. Moreover Putnam argues that in the case of a psychological, or more broadly, a 'logical' or functionally defined state, like *preferring A to B*, it is known that 'we *cannot* discover laws by virtue of which it is physically necessary that an organism prefers A to B if and only if it is in a certain physical-chemical state' since it is known to be possible for the logical states of a system to be physically realised in an indefinite number of different ways. He further points out a paradoxical consequence of identifying such states with a certain structure. Suppose, to construct my own example, we found creatures on some other planet who were sufficiently like ourselves in appearance and behaviour for us to be ready to credit them with preferences, beliefs, and so on. But then suppose that we subsequently discovered that their 'central processes' were markedly different from our own. The consequences would be that on the Identity Theory we would be *forced* to withdraw the application of our own psychological vocabulary to them – or, absurdly, to endow it with a quite different meaning. If the identity of the Identity Theory is to be taken seriously, then what belief, say, is depends on the actual physical nature of the underlying state. This does not seem acceptable. There is good reason to say that we all know perfectly well what belief is, and what it is for a person to believe something, prior to scientific discoveries about brain mechanisms, and that such discoveries add not at all to our knowledge of the nature of belief. Instead such discoveries tell us only the changes in brain conditions which occur when people come to acquire beliefs. This knowledge may, of course, be to the highest degree useful from the practical point of view of manipulating people's beliefs, whether for therapeutic or less scrupulous reasons. But what really counts, as far as belief itself is concerned, is the functionally characterised state and not the physical structure which sustains or embodies it.

Something of this is apparent even in the case of the gene and the D.N.A. molecule; genes are D.N.A. molecules but not all D.N.A.

molecules are genes: only in respect of their function of transmitting hereditary information are certain D.N.A. molecules genes. However, in this case it must certainly be allowed that the discovery of the physical nature of genes really does add to our knowledge of what genes are. And I believe that the same may well be true in the case of those concepts which the Identity Theory was originally introduced to deal with, namely, the concepts of sensation, consciousness, mental imagery; and of course, of all other concepts just *in so far as* these too involve such concepts. This may be so, I suggest, because specific biological processes really are, as a matter of fact, physically necessary to consciousness. Feigl, for one, remarks that 'inductively it is plausible that sentience requires complex organic processes'; though I suspect one would be hard put to it to specify just what such inductive evidence consists of.

One further suggestion concerning the nature of the relation between mental and cerebral states should be mentioned. The suggestion has been made (see Max Deutscher, 'Mental and Physical Properties', *The Identity Theory of Mind*, ed. C. F. Presley (Brisbane 1967) pp. 65-83) that the appropriate notion that Identity Theorists are in fact operating with is not really that of identification *with* but rather of identification *as*. Thus the first step of the Smart-Armstrong exposition provides an account of mental states, as has been previously indicated, involving an unknown or unspecified element. Smart, for example, compared a sensation-report with the report 'Someone telephoned', where the someone turns out to be the doctor. Thus the someone is initially characterised only relationally, or known by description, as 'having telephoned' but is subsequently *identified as* the doctor (compare J. M. Hinton's 'the way of identification' p. 253, below). Analogously, mental states would be initially characterised purely relationally as states having specific causes and effects, and then subsequently scientifically identified as particular cerebral processes. This way of looking at the matter does at least help to ease the objection concerning the symmetry of identification *with* statements and the related need for providing independent identifications (*as*) of both terms. But it does not, of course, help to remove the objection that the physico-chemical aspect may not be the important thing in understanding mental states, nor of the need for the satisfaction of Leibniz's Law.

VI. LEIBNIZ'S LAW

Most of the traditional objections to an identity between mental and

physical states have, in fact, hinged on this principle of Leibniz's Law, which governs the identity relation. The principle is that of the identity of indiscernibles, or rather the conjunction of this with its converse, the indiscernibility of identicals: if A is the very same item as B, then A and B must be indiscernible in the sense of possessing all their properties in common; and this for the very good reason that in such cases there is only one item in question, though one referred to by the different expressions 'A' and 'B'. However, stated in the form: 'A' can in any statement be substituted for 'B', and conversely, without changing the statement's truth-value, the principle is subject to numerous exceptions, as detailed presently; so that the Identity Theory is in no way threatened by a failure of substitutability which is of a kind that applies to *any* contingent identity statement.

In fact many of these traditional objections have been based on a failure to appreciate the possibility of contingent, as opposed to logical, identities. Once the contingent nature of the proposed identity is realised many of the traditional objections concerning the impossibility or absurdity of an identity between mental and physiological states or processes fail to hold up, though others still remain to be tackled. One group of objections which do disappear are those of an epistemological variety. For instance it has been customary to argue that mental states *could not* be physiological states because people had, and still have, knowledge of the former without possessing any knowledge of the latter. This argument, and others akin to it, no longer possesses any force. People were able to speak about genes before anything was known about D.N.A. molecules, but for all that genes *are* D.N.A. molecules. This point can be generalised to cover all of the so-called 'intensional predicates'; these include the whole range of what Russell called 'words for propositional attitudes', like 'believes that', 'hopes that', 'expects that', 'fears that'; and also modal predicates of the form 'is necessarily...' and 'is possibly...'. Such cases constitute exceptions to Leibniz's Law, if this is understood in terms of substitutability or if 'indiscernibility' is taken to include these alleged predicates, even although it remains true that if $A=B$, every property that A possesses B also possesses; for *ex hypothesi* A is B. That is, it can simply be denied that the so-called 'intensional properties' are really genuine properties, characteristics, or attributes of the thing in question at all. It may be true, for example, that Tom believes (or fears, etc.) that the Morning Star is likely to explode, without its also being true that Tom believes (or fears, etc.) that the Evening Star is likely to explode, even although, unknown to Tom, the Morning Star *is* the Evening Star. Again it is

necessarily true that the Morning Star is visible in the morning, yet not necessarily true that the Evening Star is visible in the morning, though both are the planet Venus. But this is surely because such expressions as 'being believed by Tom to be likely to explode' and 'being necessarily visible in the morning' do not specify genuine properties, not even relational properties, of Venus, or whatever object is in question. Be this as it may, the really important point is that since these apparent exceptions to Leibniz's Law apply to undisputed cases of identity, it is no ground for objection to the mind-brain identity proposal that they crop up over this.

Prima facie objections still remain, however, in the case of non-intensional properties. Thus it may be objected that beliefs are true, false, well-founded or absurd; after-images are yellow or green or hazy; but that no brain process could intelligibly be said to be any of these things. Conversely that brain processes are located within the skull, are fast or slow, of such-and-such an electro-chemical nature; but that mental states could not be described in these terms. Replies by advocates of the Identity Theory often turn on two important moves: either just insisting that such objections simply beg the question against the theory, sometimes coupled with the suggestion that additional rules could be adopted which would render descriptions not now applicable legitimate (see Smart, p. 62, below; Shaffer p. 116, below); or of so choosing the exact nature of the terms between which the identity is to hold that the objections simply fail to arise. These points are of particular importance, in the case of what is probably the most refractory objection of this type, namely, that concerning spatial location. This matter of location is crucial; both because, as proponents of the theory agree, same spatial location enters into the very meaning of the identity, and because, as previously mentioned, the mental has traditionally been regarded as essentially non-spatial – though still temporally located, unlike the case of abstract objects. A further move made in attempted solution has been to invoke the idea of what has been called 'partial location'. The suggestion here is that a person's pains, images or thoughts occur wherever he happens to be, so that if I experience a pain whilst in the reading room of the Library then that is where the experience occurs. So if, following a suggestion of Thomas Nagel (see p. 218, below), the two terms between which the identity is to hold are taken to be, on the one hand, my having a sensation, and on the other, my body's being in a certain physical state, then both will be located to the same degree; namely again, wherever I happen to be. Difficulties arise, however, when considering just how far this

partial location can be narrowed down. Does the having of the pain occur in one part of the room rather than another? Does it 'move' when I shift my position on the chair? It is, to be sure, perfectly natural to say that I experience a pain, or that a thought occurs to me, when I am at such and such a place. But whether this can be utilised in such a way as to remove objections concerning locatability is still a matter of dispute.

Another basic objection raised against the brain-process theory concerns the alleged incorrigibility of introspective or first-person reports of, at any rate, sensations, thoughts, images and intentions. The significance of this point is not perhaps primarily in regard to the satisfaction of Leibniz's Law, so much as the correlation between mental and cerebral states on which, at least on most versions, the Identity Theory is founded. If then I 'incorrigibly' report the occurrence of a severe pain in my left arm, but the brain scientist fails to detect the occurrence of the usual cerebral correlate, what is to be said? Smart's initial reply (see Paper IX) to Professor Baier, who raised this point (Paper VIII), was to the effect that since the identity he proposed was purely contingent, he simply believed that this would never in fact happen, but that if it did then this would refute his theory. However, towards the end of his reply Smart introduced another suggestion, namely, that first-person reports are never in fact incorrigible at all, on the ground that the report and what is reported constitute distinct occurrences so that, logically, it is always possible for one to occur without the other. (This line of thought is carried to extraordinary lengths in Armstrong's *A Materialist Theory of the Mind* where it is asserted, for example, that there can be *felt* pains of which one is unaware (see p. 312 of that book); this at least will surely not do.)

There are two immediate objections to this. In the first place, it seems extremely doubtful whether a sincere first-person report of an intense pain could ever become corrigible without a radical change in our concept of pain; as Baier remarks, it makes no sense to say 'I have a pain unless I am mistaken': there are no pain hallucinations. Equally it would seem bizarre to suggest that a man might be informed (as opposed to taught) that he was in pain. In the second place, if the incorrigibility of such reports is called in question it is somewhat difficult to see how the required psycho-physical correlations could ever be set up at all. So perhaps with this suggestion Smart is unwittingly sawing off the branch on which he is sitting.

VII. THE STATUS OF THE IDENTITY THEORY

Little has so far been said about the empirical side of the Identity Theory. Its proponents normally wish to maintain not only that it is free from logical objections but also that it is a reasonable and plausible hypothesis. This they aim to do in the light of current scientific knowledge, notably of supposed psycho-neurophysiological correlations (though if the theory is correct it would be more appropriate to speak of correlations between descriptions or observations). The subsistence of such correlations is, of course, accepted by other theories of mind. By, for example, the notorious doctrine of Epiphenomenalism, which asserts that while mental phenomena are caused by correlated brain events, they themselves are without causal influence either on any physical events or on other mental phenomena.

The contention then would be that these correlations are best interpreted in terms of identity and not in terms of relations between distinct occurrences, whether causally related or not. The recommendation is made primarily on grounds of simplicity and economy, of Occam's razor. Not only are superfluous entities disposed of but the subsistence of the correlations becomes explicable. Without this, these correlations would have to be just accepted as brute and inexplicable facts which, moreover, could not be fitted into the general scientific framework. Correlation 'laws' relating the intersubjectively confirmable with the non-intersubjectively confirmable would be, as Feigl expresses it, 'nomological danglers'.

However I doubt whether the choice between the Identity Theory and the rival theories of Epiphenomenalism and Psycho-physical Parallelism can correctly be regarded as an ordinary scientific issue. The crux, is, as would be generally acknowledged, that it is in principle impossible for any experimental test to resolve the dispute. The case is possibly different where the choice is between any of these theories and a theory of mind-brain interactionism. For in this case it is plausible to argue that the breaks in the chain of physical causation should be, in principle, empirically detectable. The hope would be that if this issue could be experimentally decided against Interactionism, then Epiphenomenalism and Psycho-physical Parallelism could be rejected on other grounds. This does not seem an unreasonable hope. For these doctrines involve the truly incredible view that none of our mental states can be causally effective: no thoughts, decisions, pains, or acts of attention would themselves have any causal result on behaviour. Moreover, the

very existence and the distribution of these states becomes totally inexplicable, both from a biological point of view and in terms of the above-mentioned framework of the natural sciences. The great strength of the Identity Theory is that it allows causal efficacy to mental occurrences: itches really do cause one to scratch, decisions result in appropriate action, attending to a task produces more efficient performance; and so on.

However, the case for the Identity Theory has not yet been convincingly made out. Some of the logical objections remain inadequately answered; the very intelligibility of the position is still in dispute; and, of course, its final acceptability must partly await further scientific research, including psychical research, the claims of which Armstrong describes as 'the small black cloud on the horizon of a Materialist theory of mind' (*A Materialist Theory of the Mind*, p. 364). The extension of the theory to cover all of our mental concepts is, to my mind, not at all plausible. It appears to depend on too primitive a conception of these concepts, roughly that all refer to specific states: almost a *unum nomen-unum nominatum* view. But, as a great deal of recent philosophising has shown, such a view is quite untenable. Consider, for example, the notion of promising. Many philosophers would contend that arguments, notably those of J. L. Austin, had established that this notion is primarily performative: to promise is to perform a socially institutionalised act. A person who promises is not reporting the existence of a state within himself which normally has the causal result of ensuring that the promised act is done. The very uttering of the words 'I promise' in appropriate circumstances is actually to make the promise; a person thereby *commits* himself to doing the act in question, he is not describing or reporting anything. Possibly he forgets about his promise until, at a later date, he is reminded of it, and, being a man of his word, keeps it. (There clearly need be no direct causal connection.) It may indeed be the case that such points can be accommodated within the Identity Theory, but if so, this has yet to be shown. Certainly no crude application of it to every psychological concept is likely to succeed.

What has, however, been established, is that the Identity Theory as such can no longer be lightly dismissed. Whether it ultimately manages to survive or not, what can at least be confidently claimed is that, at the present time, it provides one very fruitful focus of interest for discussions of problems in the philosophy of mind.