

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

# XXI Physicalism

Thomas Nagel

## I

It is the purpose of this paper to examine the reasons for believing that physicalism cannot possibly be true.<sup>1</sup> I mean by physicalism the thesis that a person, with all his psychological attributes, is nothing over and above his body, with all its physical attributes. The various theories which make this claim may be classified according to the identities which they allege between the mental and the physical.<sup>2</sup> These identities may be illustrated by the standard example of a quart of water which is identical with a collection of molecules, each containing two atoms of hydrogen and one of oxygen.

All states of the water are states of that collection of molecules: for the water to be in a particular bottle is for those molecules to be arranged in that bottle; for the water to be frozen is for the molecules to be arranged in a space lattice, with strong intermolecular attractive force and relatively weak individual vibratory motion; for the water to be boiling is for the molecules to have a kinetic energy sufficient to produce a vapour pressure equal to the atmospheric pressure; and so forth. In addition to general identities like these, there will be particular ones.<sup>3</sup> One such is the identity between an individual splash of the water and a particular sudden displacement of certain of the molecules – an identity which does not imply that a splash is always identical with that particular type of displacement. In all of these cases we can say something like the following: that the water's splashing is not anything over and above the displacement of those molecules; they are the same occurrence.

<sup>1</sup> An earlier version of this paper was read at the Pacific Division American Philosophical Association meetings in Seattle, 5 September 1964.

<sup>2</sup> I shall not consider behaviourism or reductionism of any kind.

<sup>3</sup> Any identity both of whose terms are universal in form will be called general, even if their specification involves reference to particulars. Thus, 'Water is  $H_2O$ ', 'For water to be frozen is for its molecules to be in condition *F*', and 'For *this* water to be frozen is for its molecules to be in condition *F*' are all general identities. On the other hand, 'This water's (now) being frozen is its molecules' being in condition *F*' is a particular identity.

It is not clear whether every physicalist theory must assert the identity of each person with his body, nor is the connection between this identity and that of psychological with physical states easy to describe. Still, we can specify a range of possible views in terms of the latter relation alone. (1) An implausibly strong physicalism might assert the existence of a general identity between each psychological condition and a physical counterpart. (2) A weaker view would assert some general identities, particularly on the level of sensation, and particular identities for everything that remains. (3) A still weaker view might not require that a physical condition be found identical even in the particular case with every psychological condition, especially if it were an intensional one. (4) The weakest conceivable view would not even assert any particular identities, but of course it is unclear what other assertion by such a theory about the relation between mental and physical conditions might amount to a contention of physicalism.

I am inclined to believe that some weak physicalist theory of the third type is true, and that any plausible physicalism will include some state and event identities, both particular and general. Even a weak view, therefore, must be defended against objections to the possibility of identifying *any* psychological condition with a physical one. It is with such general objections that we shall be occupied.

I shall contend that they fail as objections to physicalism, but I shall also contend that they fail to express properly the real source of unhappiness with that position. This conclusion is drawn largely from my own case. I have always found physicalism extremely repellent. Despite my current belief that the thesis is true, this reaction persists, having survived the refutation of those common objections to physicalism which I once thought expressed it. Its source must therefore lie elsewhere, and I shall make a suggestion about that later.<sup>4</sup> First, however, it will be necessary to show why the standard objections fail, and what kind of identity can hold between mental and physical phenomena.

## II

Since Smart refuted them, it has presumably become unnecessary to discuss those objections which rest on the confusion between

<sup>4</sup> In sec. v; of the other sections, II attempts to rebut some standard objections, and III contains a general discussion of identity whose results are applied to physicalism in IV.

identity of meaning and identity in fact.<sup>5</sup> We may concentrate rather on two types of objection which seem still to be current.

The first is that physicalism violates Leibniz's law, which requires that if two things are identical they have all their non-intensional and non-modal properties in common. It is objected that sensory impressions, pains, thoughts, and so forth have various properties which brain states lack, and vice versa. I shall eventually propose a modification of Leibniz's law, since I do not believe that in its strict form it governs the relation asserted by the identity thesis. At this point, however, the thesis may be defended without resorting to such methods, through a somewhat altered version of a device employed by Smart, and earlier by U. T. Place.<sup>6</sup>

Instead of identifying thoughts, sensations, after-images, and so forth with brain processes, I propose to identify a person's having the sensation with his body's being in a physical state or undergoing a physical process. Notice that both terms of this identity are of the same logical type, namely (to put it in neutral terminology) a subject's possessing a certain attribute. The subjects are the person and his body (not his brain), and the attributes are psychological conditions, happenings, and so forth, and physical ones. The psychological term of the identity must be the person's having a pain in his shin rather than the pain itself, because although it is undeniable that pains exist and people have them, it is also clear that this describes a condition of one entity, the person, rather than a relation between two entities, a person and a pain. For pains to exist is for people to have them. This seems to me perfectly obvious, despite the innocent suggestions of our language to the contrary.

So we may regard the ascription of properties to a sensation simply as part of the specification of a psychological state's being ascribed to the person. When we assert that a person has a sensation of a certain description *B*, this is not to be taken as asserting that there exist an *x* and a *y* such that *x* is a person and *y* is a sensation and *B*(*y*), and *x* has *y*. Rather we are to take it as asserting the existence of only one thing, *x*, such that *x* is a person, and moreover *C*(*x*), where *C* is the attribute 'has a sensation of description *B*'. The specification of this attribute is accomplished in part by the ascription of properties to the sensation; but this is merely part of

<sup>5</sup> See Paper III; also Paper XVI, and his book *Philosophy and Scientific Realism* (London 1963), for fuller discussion of the identity thesis.

<sup>6</sup> Place, pp. 49–50, above; for Smart, see p. 61, above. My formulation of the identity differs from Smart's, and I do not accept his psychological reductionism.

the ascription of that psychological state to the person. This position seems to me attractive independently of physicalism, and it can be extended to psychological states and events other than sensations. Any ascription of properties to them is to be taken simply as part of the ascription of other attributes to the person who has them – as *specifying* those attributes.

I deviate from Smart in making the physical side of the identity a condition of the body rather than a condition of the brain,<sup>7</sup> because it seems to me doubtful that anything without a body of some conventional sort could be the subject of psychological states.<sup>8</sup> I do not mean to imply that the presence of a particular sensation need depend on the condition of any part of one's body outside of the brain. Making the physical term of the identity a bodily rather than a brain state merely implies that the brain is *in* a body. To identify the person's having a pain with the brain's being in state *X* rather than with the body's containing a brain in state *X* would imply, on the other hand, that if the individual's brain could have been in that state while the rest of his body was destroyed, he would still have been in the corresponding psychological state.

Given that the terms of the identity are as specified, nothing obliges us to identify a sensation or a pain or a thought with anything physical, and this disposes of numerous objections. For although I may have a visual sense impression whose attributes of form and colour correspond closely to those which characterise the 'Mona Lisa', my *having* the sense impression does not possess those attributes, and it is therefore no cause for worry that nothing in my brain looks like the 'Mona Lisa'. Given our specification of the psychological side of the identity, the demands on the physical side are considerably lessened. The physical equivalents of auditory impressions may be silent, those of olfactory impressions odourless, and so forth.

Most important, we can be rid of the stubbornest objection of this type, that having to do with location.<sup>9</sup> Brain processes are located in the brain, but a pain may be located in the shin and a thought has

<sup>7</sup> One might alternatively make it a physical condition of the *person*, so that the two identified attributes would be guaranteed the same subject. I cannot say how such a change would affect the argument.

<sup>8</sup> Cf. Malcolm, pp. 179–80, above.

<sup>9</sup> Malcolm, pp. 174–5, above. See also Shaffer, Paper X. Shaffer thinks the difficulty can be got over, but that this depends on the possibility of a *change* in our concept of mental states, which would make it meaningful to assign them locations.

no location at all. But if the two sides of the identity are not a sensation and a brain process, but my *having* a certain sensation or thought and my body's *being* in a certain physical state, then they will both be going on in the same place – namely, wherever I (and my body) happen to be. It is important that the physical side of the identity is not a brain process, but rather my *body's* being in that state which may be specified as 'having the relevant process going on in its brain'. *That* state is not located in the brain; it has been located as precisely as it can be when we have been told the precise location of that of which it is a state – namely, my body. The same is true of my having a sensation: that is going on wherever I happen to be at the time, and its location cannot be specified more precisely than mine can. (That is, even if a pain is located in my right shin, I am *having* that pain in my office at the university.) The location of bodily sensations is a very different thing from the location of warts. It is phenomenal location, and is best regarded as one feature of a psychological attribute possessed by the *whole* person rather than as the spatial location of an event going on in a part of him.

The other type of objection which I shall discuss is that physicalism fails to account for the privacy or subjectivity of mental phenomena. This complaint, while very important, is difficult to state precisely.

There is a trivial sense in which a psychological state is private to its possessor, namely, that since it is his, it cannot be anyone else's. This is just as true of haircuts or, for that matter, of physiological conditions. Its triviality becomes clear when we regard thoughts and sensations as conditions of the person rather than as things to which the person is related. When we see that what is described as though it were a relation between two things is really a condition of one thing, it is not surprising that only one person can stand in the said relation to a given sensation or feeling. In this sense, bodily states are just as private to their possessor as the mental states with which they may be equated.

The private-access objection is sometimes expressed epistemologically. The privacy of haircuts is uninteresting because there is lacking in that case a special connection between possession and knowledge which is felt to be present in the case of pains. Consider the following statement of the privacy objection.<sup>10</sup> 'When I am in a

<sup>10</sup> See Baier, Paper VIII, and Smart, Paper IX. This is regarded as a serious difficulty by Smart and other defenders of physicalism. See Armstrong, 'Is Introspective Knowledge Incorrigible?', *Philosophical Review*, LXXII (1963) pp. 418–19. On the other hand, Hilary Putnam has argued

psychological state – for example, when I have a certain sensation – it is logically impossible that I should fail to know that I am in that state. This, however, is not true of any bodily state. Therefore no psychological state is identical with any bodily state.' As it happens, I believe that the first clause of this objection – namely, the incorrigibility thesis – is false, but I do not have to base my attack on that contention, for even if the incorrigibility thesis were true it would not rule out physicalism.

If state *x* is identical with state *y* it does not follow by Leibniz's law that if I know I am in state *x* then I know I am in state *y*, since the context is intensional. Therefore neither does it follow from 'If I am in state *x* then I know I am in state *x*' that if I am in state *y* I know I am in state *y*. All that follows is that if I am in state *y* I know I am in state *x*. Moreover, this connection will not be a necessary one, since only one of the premises – the incorrigibility thesis – is necessary. The other premise – that *x* is identical with *y* – is contingent, making the consequence contingent.<sup>11</sup>

There may be more to the special-access objection than this, but I have not yet encountered a version of it which succeeds. We shall later discuss a somewhat different interpretation of the claim that mental states are subjective.

### III

Let us now consider the nature of the identity which physicalism asserts. Events, states of affairs, conditions, psychological and otherwise, may be identical in a perfectly straightforward sense which conforms to Leibniz's law as strictly as does the identity between,

that all the problems about privacy and special access which can be raised about persons can be raised about machines as well. See his paper 'Minds and Machines', *Dimensions of Mind*, pp. 148–79.

<sup>11</sup> It is worth noting that if two mental states are necessarily connected, this connection must be mirrored on the level of the physical states with which we identify them. Although the connection between the physical states need not be a logically necessary one, that would be a desirable feature in a physicalistic theory, and it seems in fact to be present in the example of water and molecules: the water's being frozen necessarily includes its being cold, and the specification of the molecular state which is its being frozen entails that the molecules will have a low average kinetic energy – which is in fact the same thing as the water's being cold.

say, the only horse in Berkeley and the largest mammal in Berkeley. Such identities between events may be due to the identity of two things referred to in their descriptions – for example, my being kicked by the only horse in Berkeley and my being kicked by the largest mammal in Berkeley – or they may not – for example, the sinking of the *Titanic* and the largest marine disaster ever to occur in peacetime. Whether they hold between things, events, or conditions, I shall refer to them as *strict* identities.

We are interested, however, in identities of a different type – between psychological and physical events, or between the boiling of water and the activity of molecules. I shall call these theoretical identities<sup>12</sup> and shall concentrate for the moment on their application to events and attributes rather than to things, although they hold between things as well. It is a weaker relation than strict identity, and common possession of causal and conditional attributes is crucial for its establishment.<sup>13</sup> Strict identities are likely to be established in other ways, and we can infer the sameness of all causal and conditional attributes. Thus, if being kicked by the only horse in Berkeley gave me a broken leg, then being kicked by the largest mammal in Berkeley had the same effect, given that they are the same creature; and if it is the case that I should not have been kicked by the only horse in Berkeley if I had stayed in my office that afternoon, then it follows that if I had stayed in my office I should not have been kicked by the largest mammal in Berkeley.

But if we lack grounds such as these, we must establish sameness of conditional attributes independently, and this depends on the discovery of general laws from which the particular conditionals follow. Our grounds for believing that a particular quart of water's

<sup>12</sup> Following Hilary Putnam, in *Dimensions of Mind*, who says that the 'is' in question is that of theoretical identification. The word 'identity' by itself is actually too strong for such cases, but I shall adopt it for the sake of convenience.

<sup>13</sup> An attribute, for our purposes, is signified by any sentence-frame containing one free variable (in however many occurrences) where this may be a variable ranging over objects, events, and so forth. (One gets a particular instance of an attribute by plugging in an appropriate particular for the variable and converting to gerundival form.) Thus all three of the following are attributes: '... is boiling', '... will stop boiling if the kettle is taken off the fire', and '... will stop if the kettle is taken off the fire'. A particular quart of water has the second of these attributes if and only if that water's boiling has the third, where this can be described as the possession of the third attribute by a particular instance of the first.

boiling is the same event as a collection of molecules' behaving in a certain way are whatever grounds we may have for believing that all the causes and effects of one event are also causes and effects of the other, and that all true statements about conditions under which the one event would not have occurred, or about what would have happened if it had not, or about what would happen if it continued, and so forth, are also true of the other.

This is clearly more than mere constant conjunction; it is a fairly strong requirement for identity. Nevertheless it is weaker than the standard version of Leibniz's law in that it does not require possession by each term of *all* the attributes of the other. It does not require that the complex molecular event which we may identify with my being kicked by the only horse in Berkeley be independently characterisable as ridiculous – for example, on the grounds that the latter event was ridiculous and if the former cannot be said to be ridiculous, it lacks an attribute which the latter possesses. There are some attributes from the common possession of which the identity follows, and others which either do not matter or which we cannot decide whether to ascribe to one of the terms without first deciding whether the identity in question holds.

To make this precise, I shall introduce the notion of independent ascribability. There are certain attributes such as being hot or cold, or boiling or offensive, which cannot significantly be ascribed to a collection of molecules *per se*. It may be that such attributes *can* be ascribed to a collection of molecules, but such ascription is dependent for its significance on their primary ascription to something of a different kind, like a body of water or a person, with which the molecules are identical. Such attributes, I shall say, are not independently ascribable to the molecules, though they may be dependently ascribable. Similarly, the property of having eighty-three trillion members is not independently ascribable to a quantity of water, though it may be possessed by a collection of H<sub>2</sub>O molecules. Nevertheless, there is in such cases a class of attributes which are independently ascribable to both terms, and the condition for theoretical identity may be stated as follows: that the two terms should possess or lack in common all those attributes which can be independently ascribed to each of them individually – with the qualification that nothing is by this criterion to be identical with two things which are by the same criterion distinct.<sup>14</sup> Actually this

<sup>14</sup> The qualification takes care of such possibly problematic claims as that I am the square root of 2, for although it may be that we share all attributes which can be independently ascribed to each of us, I also share those



will serve as a condition for identity in general; a strict identity will simply be one between terms sufficiently similar in type to allow independent ascription to both terms of *all* the same attributes, and will include such cases as the sinking of the *Titanic* being the largest marine disaster ever to occur in peacetime, or the Morning Star being the Evening Star. The identities I have characterised as theoretical hold across categories of description sufficiently different to prohibit independent ascription to both terms of all the same attributes, although, as I have observed, such ascriptions may be meaningful as *consequences* of the identity.

The question naturally arises, to what extent do particular theoretical identities depend on corresponding general ones? In the examples I have given concerning the case of water, the dependence is obvious. There the particular identities have simply been instances of general ones, which are consequences of the same theory that accounts for the common possession of relevant attributes in the particular cases. Now there is a technical sense in which every particular theoretical identity must be an instance of a general identity, but not all cases need be like that of water. Although it is essential that particular identities must follow from general laws or a general theory, this does not prevent us from distinguishing between cases in which, for example, the molecular counterpart of a macroscopic phenomenon is always the same, and those in which it varies from instance to instance. The common possession of conditional attributes can follow for a particular case from general laws, without its being true that there is a general correlation between macroscopic and microscopic phenomena of that type. For example, it may at the same time follow from general laws that types of microscopic phenomena other than the one present in this case would also share the requisite conditional properties.

The technical sense in which even in such cases the particular identity must be an instance of a general one is that it must be regarded as an instance of the identity between the macroscopic phenomenon and the disjunction of all those microscopic phenomena which are associated with it in the manner described, via general laws. For suppose we have a type of macroscopic phenomenon *A* and two types of microscopic phenomena *B* and *C* so associated with it. Suppose on one occasion particular cases of *A* and *B* are occurring at the same place and time, and so forth, and suppose it

attributes with the square root of 3, whose attributes clearly contradict those of the square root of 2.

is asserted that since it follows from general laws that they also have all their conditional attributes in common, *A* is in this case identical in the specified sense with *B*. They do not, however, have in common the conditional attribute *F(X)*, defined as follows: 'If *C* and not *B*, then *X*.' That is, *F(A)* but not *F(B)*. Therefore, we must identify the occurrence of *A* even in this case with the occurrence of the disjunction *B* or *C*. This does not prevent us, however, from introducing as a subsidiary sense of identity for particular cases that in which *A* is *B* because the disjunction *B* or *C* which is properly identical with *A* is in fact satisfied by *B*. There is of course a range of cases between the two kinds described, cases in which the disjuncts in the general identity consist of conjunctions which overlap to a greater or lesser degree, and this complicates the matter considerably.<sup>15</sup> Nevertheless we can, despite the technicality, differentiate roughly between particular identities which are in a narrow sense instances of general identities and those which are not – that is, which are instances only of radically disjunctive general identities. Henceforth when I refer to general identities I shall be excluding the latter.

I have concentrated on identities between states, events, and attributes because it is in such terms that physicalism is usually conceived, but if it is also part of physicalism to hold that people are their bodies it becomes appropriate to inquire into the relation between the theoretical identity of things and the theoretical identity of their attributes. Unfortunately, I do not have a general answer to this question. The case of strict identity presents no problem, for there every attribute of one term is strictly identical with the corresponding attribute of the other; and in our standard example of theoretical identity, each attribute of the water seems to be theoretic-

<sup>15</sup> A fuller treatment would have to include a discussion of the non-symmetrical relation '...consists of...' which is distinct from identity. A macroscopic event (the freezing of some water, for example) may be identical with a microscopic event *A* described in general terms (average kinetic energy, spatial ordering, and the like) while at the same time consisting of a very specific collection *B* of microscopic events with which it is not identical, since if one of them (the motion of a particular molecule) had been different, that particular complex of microscopic events would not have occurred though both *A* and the macroscopic event would have. (Presumably in such cases the occurrence of *B* entails the occurrence of *A*, but more than that needs to be said about the relation between the two.) The same concept applies to the relation between World War II and the immense collection of actions and events of which it consisted, or that between the Eiffel Tower and the girders and rivets which make it up.

cally identical with some attribute of the molecules, but not vice versa. This may be one (asymmetrical) condition for the theoretical identity of things. It is not clear, however, whether the identity of things must always be so closely tied to the identity of their attributes. For example, it might be that everything we could explain in terms of the water and its attributes could be explained in terms of the batch of molecules and their attributes, but that the two systems of explanation were so different in structure that it would be impossible to find a single attribute of the molecules which explained all and only those things explained by a particular attribute of the water.

Whether or not this is true of water, the possibility has obvious relevance to physicalism. One might be able to define a weak criterion of theoretical identity which would be satisfied in such a case, and this might in turn give sense to an identification of persons with their bodies which did not depend on the discovery of a single physical counterpart for every psychological event or condition. I shall, however, forgo an investigation of the subject; this general discussion of identity must remain somewhat programmatic.

#### IV

It provides us with some things to say, however, about the thesis of physicalism. First, the grounds for accepting it will come from increased knowledge of (a) the explanation of mental events and (b) the physiological explanation of happenings which those mental events in turn explain. Second, in view of the condition of independent ascribability, physicalism need not be threatened by the difficulty that although anger may be, for example, justified, it makes no sense to say this of a physical state with which we may identify it. Third, it does not require general identities at every level: that is, there need not be, either for all persons or even for each individual, a specific physical state of the body which is *in general* identical with intending to vote Republican at the next election, or having a stomach-ache, in order that physicalism be true. It seems likely that there will be general identities of a rough kind for non-intensional states, such as having particular sensations or sensory impressions, since the physical causes of these are fairly uniform. But one can be practically certain that intensional mental states, even if in each particular case they are identical with some physical state, will not

have general physical counterparts, because both the causes and the effects of a given belief or desire or intention are extremely various on different occasions even for the same individual, let alone for different persons. One might as easily hope to find a general equivalent, in molecular terms, of a building's collapsing or a bridge's being unsafe – yet each instance of such an event or circumstance is identical with some microscopic phenomenon.

The relation of intensional mental states to physical states may be even more involved than this. For one thing, if it should be the case that they are dispositional in a classical sense, then physicalism requires only that the events and states to which they are the dispositions be identical with physical events and states. It does not require that they be identical with any additional independent physical state, existing even when the disposition is not being exercised. (In fact, I do not believe that dispositions operate according to the classical Rylean model, and this will affect still further the way in which the identity thesis applies to dispositional mental states; but this is not the place for a discussion of that issue.)

There is still another point: many intensional predicates do not just ascribe a condition to the person himself but have implications about the rest of the world and his relation to it. Physicalism will of course not require that these be identical simply with states of the person's body, narrowly conceived. An obvious case is that of knowledge, which implies not only the truth of what is known but also a special relation between this and the knower. Intentions, thoughts, and desires may also imply a context, a relation with things outside the person. The thesis that all states of a person are states of his body therefore requires a liberal conception of what constitutes a state – one which will admit relational attributes. This is not peculiar to mental states: it is characteristic of intensional attributes wherever they occur. That a sign says that fishing is forbidden does not consist simply in its having a certain geometrically describable distribution of black paint on its surface; yet we are not tempted here to deny that the sign is a piece of wood with paint on it, or to postulate a non-corporeal substance which is the subject of the sign's intensional attributes.

Even with all these qualifications, however, it may be too much to expect a specific physical counterpart for each particular psychological phenomenon. Thus, although it may be the case that what explains and is explained by a particular sensation can also explain and be explained by a particular neurological condition, it may also be that this is not precisely true of an intention, but rather that



the various connections which we draw between causes and effects via the intention can be accounted for in terms of many different physical conditions, some of which also account for connections which in psychological discourse we draw via states other than the intention, and no subset of which, proper or improper, accounts for all and only those connections which the intention explains. For this reason a thorough-going physicalism might have to fall back on a criterion for identity between things not dependent on the identity of their attributes – a criterion of the sort envisaged at the end of the previous section.

Obviously any physicalistic *theory*, as opposed to the bare philosophical thesis of physicalism, will be exceedingly complex. We are nowhere near a physical theory of how human beings work, and I do not know whether the empirical evidence currently available indicates that we may expect with the advance of neurology to uncover one. My concern here has been only to refute the philosophical position that mental-physical identity is *impossible*, and that *no* amount of further information could constitute evidence for it.

## V

Even if what might be called the standard objections have been answered, however, I believe that there remains another source for the philosophical conviction that physicalism is impossible. It expresses itself crudely as the feeling that there is a fundamental distinction between the subjective and the objective which cannot be bridged. Objections having to do with privacy and special access represent attempts to express it, but they fail to do so, for it remains when they have been defeated. The feeling is that I (and hence any 'I') cannot be a mere physical object, because I possess my mental states: I am their *subject*, in a way in which no physical object can possibly be the subject of its attributes. I have a type of internality which physical things lack; so in addition to the connection which all my mental states do admittedly have with my body, they are also mine – that is, they have a particular *self* as subject, rather than merely being attributes of an object. Since any mental state must have a self as subject, it cannot be identical with a mere attribute of some object like a body, and the self which is its subject cannot therefore be a body.

Why should it be thought that for *me* to have a certain sensation

– to be in a certain mental state – cannot consist merely in a physical object's being in some state, having some attribute? One might put it as follows. States of my body, physical states, are, admittedly, physical states of me, but this is not simply because they are states of that body but because in addition it is my body. And its being my body consists in its having a certain relation, perhaps a causal one, to the subject of my mental states. This leads naturally to the conclusion that I, the subject of my mental states, am something else – perhaps a mental substance. My physical states are only derivatively mine, since they are states of a body which is mine in virtue of being related in the appropriate way to my psychological states. But this is possible only if those psychological states are mine in an original, and not merely derivative, sense; therefore *their* subject cannot be the body which is derivatively mine. The feeling that physicalism leaves out of account the essential subjectivity of psychological states is the feeling that nowhere in the description of the state of a human body could there be room for a physical equivalent of the fact that *I* (or any self), and not just that body, am the subject of those states.

This, so far as I can see, is the source of my uneasiness about physicalism. Unfortunately, whatever its merits, it is no more an argument against physicalism than against most other theories of mind, including dualism, and it therefore provides us with no more reason for rejecting the former in favour of the latter than do the standard objections already discussed. It can be shown that if we follow out this type of argument, it will provide us with equally strong reasons for rejecting any view which identifies the subject of psychological states with a substance and construes the states as attributes of that substance. A non-corporeal substance seems safe only because, in retreating from the physical substance as a candidate for the self, we are so much occupied with finding a subject whose states are originally, and not just derivatively, mine – one to which the physical body can be related in a way which explains how *it* can be mine – that we simply postulate such a subject without asking ourselves whether the same objections will not apply to it as well: whether indeed any substance can possibly meet the requirement that its states be *underderivatively* mine.

The problem can be shown to be general in the following way: consider everything that can be said about the world without employing any token-reflexive expressions.<sup>10</sup> This will include the

<sup>10</sup> I.e. expression *functioning* as token reflexives. Such words of course lose this function in quotation and in certain cases of *oratio* (or *cogitatio obliqua*: e.g. 'John Smith thinks that he is Napoleon.')

description of all its physical contents and their states, activities, and attributes. It will also include a description of all the persons in the world and their histories, memories, thoughts, sensations, perceptions, intentions, and so forth. I can thus describe without token reflexives the entire world and everything that is happening in it – and this will include a description of Thomas Nagel and what he is thinking and feeling. But there seems to remain one thing which I cannot say in this fashion – namely, which of the various persons in the world *I* am. Even when everything that can be said in the specified manner has been said, and the world has in a sense been completely described, there seems to remain one fact which has not been expressed, and that is the fact that I am Thomas Nagel. This is not, of course, the fact ordinarily conveyed by those words, when they are used to inform someone else who the *speaker* is – for that could easily be expressed otherwise. It is rather the fact that *I* am the subject of *these* experiences; this body is my body; the subject or centre of my world is this person, Thomas Nagel.

Now it follows from this not only that a sensation's being mine cannot consist simply in its being an attribute of a particular body; it follows also that it cannot consist in the sensation's being an attribute of a particular soul which is joined to that body; for nothing in the specification of that soul will determine that *it* is mine, that I am *that* person. So long as we construe psychological states as attributes of a substance, no matter what substance we pick, it can be thrown, along with the body, into the 'objective' world; its states and its relation to a particular body can be described completely without touching upon the fact that I am that person.<sup>17</sup> It turns out therefore that, given the requirements which led us to reject physicalism, the quest for the self, for a substance which *is* me and whose possession of a psychological attribute will *be* its being mine, is a quest for something which could not exist. The only possible conclusion is that the self is not a substance, and that the special kind of possession which characterises the relation between me and my psychological states cannot be represented as the possession of certain attributes by a subject, no matter what that subject may be. The subjectivity of the true psychological subject is of a different kind from that of the mere subject of attributes. And if I am to extend this to cases other than my own, I must conclude that for no person is it the case that his having a particular sensation consists in some occupant of the world having a particular attribute or being in a certain state.

<sup>17</sup> Cf. Wittgenstein, *Tractatus*, 5.64. (London 1922).

I shall not discuss the reasons for rejecting this position. My attitude towards it is precisely the reverse of my attitude toward physicalism, which repels me although I am persuaded of its truth. The two are of course related, since what bothers me about physicalism is the thought that I cannot be a mere physical object, cannot in fact be anything *in* the world at all, and that my sensations and so forth cannot be simply the attributes of some substance.

But if we reject this view (as it seems likely that we must) and accept the alternative that a person is something in the world and that his mental states are states of that thing, then there is no *a priori* reason why it should not turn out to be a physical body and those states physical states. We are thus freed to investigate the possibility, and to seek the kind of understanding, of psychological states which will enable us to formulate specific physicalistic theories as neurology progresses.

#### POSTSCRIPT 1968

I now believe that theoretical identity is not distinct from strict identity, and that the device by which I formerly defined theoretical identity can be used to explain how Leibniz's law is satisfied by identities whose terms are of disparate types.

Suppose boiling is independently ascribable to a quart of water but not to the molecules which compose it. Nevertheless we can say that the molecules are boiling if they bear a certain relation to the water, and the water is boiling. The relation in question, call it *R*, is simply that which I formerly described as theoretical identity. It holds between *a* and *b* if (i) they possess or lack in common all those attributes which can be independently ascribed to each of them individually (call this relation *S*), and (ii) neither *a* nor *b* bears relation *S* to any third term which does not bear relation *S* to the other.<sup>18</sup> Let *F* range over non-intensional and non-modal attributes, and let us symbolise the modal statement '*F* is independently ascribable (truly or falsely) to *a*' as *I(F,a)*. Then

$$(1) S(a,b) \equiv df \langle F(I(F,a).I(F,b)). \supset .F(a) \equiv F(b) \rangle$$

$$(2) R(a,b) \equiv df \langle (i) S(a,b) \ \& \ (ii) \text{ a true statement results whenever a name or definite description is substituted for 'x' in the schema } S(a,x) \equiv S(b,x) \rangle$$

<sup>18</sup> Condition (ii) is added for the reason cited in footnote 14 of my 1965 paper.

I claim that a true statement results whenever names or definite descriptions are substituted for 'x' and 'y' in the following schema:

$$(3) (F)(I(F,x).F(x).R(x,y). \supset F(y))$$

If this is correct, then when *a* and *b* are related by *R* they will share all the attributes independently ascribed to either of them. By Leibniz's law, therefore,

$$(4) R(a,b) \equiv a = b.$$

## XXII Mind-body identity, a side issue?

Charles Taylor

In recent years, a number of philosophers have argued in favour of materialism in the form of an identity thesis – that is, a thesis to the effect that mental events are identical with certain physiological events.<sup>1</sup> Some have argued directly in favour of a thesis of this kind;<sup>2</sup> others have tried to show that it at least escapes the charge of logical incoherence which has been levelled against it;<sup>3</sup> the implication is that, if this charge can be set aside, the thesis is quite plausible.

It is more rarely questioned, however, just how plausible the ground is on which the thesis rests, and whether it really amounts to an affirmation of materialism. I should like to raise both these questions.

The attempt is made by identity theorists to show that an event considered typical of the mental, like a sensation or the having of an after-image, can be considered without contradiction or incoherence to be identical with a process in the brain or some physiological process in the body as a whole, and that, if further research in neurophysiology, biochemistry, and so on turns out as it is plausible to expect it will, then the identity thesis will impose itself as beyond further objection.

The argument therefore turns on the question whether, granted certain results expected from physiology, biochemistry, and so forth, we can speak of an identity between, say, the having of after-images and the occurrence of brain processes, or the having of sensations and the occurrence of physiological processes in the body as a whole. The argument therefore turns on whether these results will permit us to say that, according to the generally accepted criterion of

<sup>1</sup> Cf., among others, Place, Paper II; Feigl, "The "Mental" and the "Physical", *Minnesota Studies in the Philosophy of Science*, II, pp 370–497; Smart, *Philosophy and Scientific Realism* (London 1963).

<sup>2</sup> E.g. Smart, *Philosophy and Scientific Realism* and Place, Paper II.

<sup>3</sup> E.g. Rorty, Paper XX; and Nagel, Paper XXI.