

JAEGWON KIM

MATERIALISM AND THE CRITERIA OF THE MENTAL

I. INTRODUCTION

Even the materialist who asserts that all mental events are physical events in the brain must begin with some reasonably clear and determinate concept of what a mental event is. How else would anyone, including the materialist himself, know what he is talking about? Thus, J. J. C. Smart begins his exposition and defense of the identity theory of mind with what seems at least like an informal characterization of his subject matter:

Suppose that I report that I am having an orange-yellow roundish after-image.... There seems to be some element of 'pure inner experience' which is being reported, and to which only I have direct access. You can observe my behaviour, but only I can be aware of my own after-image or my own pain....¹

And U. T. Place, another prominent proponent of the identity theory, follows a similar route:

For our present purposes, however, I shall assume... that statements about pains and twinges, about how things look, sound, and feel, about things dreamed of or pictures in the mind's eye, are statements referring to events and processes which are in some sense private or internal to the individual of whom they are predicated.²

These characterizations of mental events are similar in that they both make use of such essentially epistemic concepts as privacy and direct accessibility; and when Smart and Place claim that mental events are brain events, they must be understood as claiming that events with these properties are brain events. And unless the existence of mental events is denied, which is not likely, these materialists are committed to the thesis, which might seem absurd, that certain physico-chemical events inside the skull have the properties of privacy and direct accessibility.

The problem of course is that once a certain property is designated as the defining characteristic of mental events, this definitional act may by itself exclude certain forms of materialism (and certain forms of idealism, too), such as the identity theory. Clearly, if we define 'mental event' by a certain property (e.g., privileged access in some sense) and then define

'physical event' as those events lacking it, or conversely define 'physical event' by a certain property (e.g., having a location in space) and 'mental event' as those lacking it, then the identification of the mental with the physical becomes hopeless; it no longer is an open possibility. To avoid this the mind-body theorist may begin with two logically independent criteria, one for marking 'mental' from 'nonmental' and the other for marking 'physical' from 'non-physical'; he could then go on to ponder the question of identity of the mental with the physical.

The purpose of this paper is to examine the problem of characterizing the mental and the physical within a materialist framework – more specifically, the framework of the identity theory. In particular, we shall look into the three major ways of drawing the mental-physical distinction and their implications for the identity theory; the three I have in mind are, as we shall call them, 'the intentionality criterion', 'the nonspatiality criterion', and 'the epistemic criterion'. Other criteria have been used to demarcate the mental from the physical, such as 'purposiveness', 'unity', and 'emergence'; these, however, will not be discussed since they seem much less central to the concept of mental current in philosophical discussion today, and they do not seem to have been formulated as clearly or as much in detail as the three to be discussed. Also, we shall, as a rule, not be concerned with the details of the particular formulations that have been given of the three criteria except those that directly affect the issue of materialism; whether these criteria can ultimately be made to work depends of course on those details.

Before we enter into the discussion of substantive issues, we should note the fact that the materialist need not provide logically independent criteria for the mental and the physical. Whether he must depends on the categorial nature of the entities which are characterized as mental or physical; in particular, by drawing a mental-physical distinction with respect to linguistic expressions, the materialist may formulate a linguistic version of the identity theory to the effect that whatever is 'described' by a mental expression is also describable by a physical expression, or that a 'complete description' of the world can be given in physical language.³ Since he is not saying that mental expressions are identical with physical expressions, there is no need for independent definitions of 'mental expression' and 'physical expression'. Thus, two different approaches toward a characterization of the mental and the physical

can be distinguished: the *linguistic approach* attempts to characterize linguistic units (predicates, sentences, languages, etc.) as mental or physical, and the *ontological approach* deals directly with nonlinguistic entities such as events, states, properties, and facts, and attempt to pick out those that are mental and those that are physical. As we shall see, the linguistic approach is at least as common in the literature as the ontological approach. In the final section of this paper, we shall remark on some aspects of these two approaches.

II. THE CRITERION OF INTENTIONALITY

In a well-known passage in his *Psychologie vom empirischen Standpunkt*, Brentano claimed that mental phenomena can be characterized by "the intentional inexistence of an object" or "the reference to a content, a direction upon an object". He writes:

This intentional inexistence is exclusively characteristic of mental phenomena. No physical phenomenon manifests anything similar. Consequently, we can define mental phenomena by saying that they are such phenomena as include an object intentionally within themselves.⁴

Brentano's approach is ontological; he is stating a defining characteristic for mental *phenomena*, not mentalistic expressions or sentences. Moreover, Brentano's observation that "no physical phenomenon manifests anything similar" suggests a definition of 'physical' as 'nonmental'. And under these definitions of 'mental' and 'physical', the identity theory would be ruled out as a logical possibility, unless of course there were no mental events, making the theory a vacuous truth.

Modern defenders of Brentano's thesis of intentionality, however, have taken a linguistic approach; notably, Roderick M. Chisholm, who is largely responsible for the current interest in intentionality, has given it many elaborate linguistic formulations which avoid entanglements with 'intentionally inexistent objects' and other metaphysical issues immaterial to the problem of characterizing the mental. According to Chisholm, intentionality is to be taken as a logico-grammatical property of expressions, principally sentences, and what is to be characterized by the help of this concept is not the concept of mental phenomenon but rather that of *mental* or *psychological sentence*. Our primary concern here is not with the specific formulations of the intentionality criterion of

psychological sentence that Chisholm has proposed; rather, our interest lies in the thesis (Chisholm calls it “Brentano’s thesis”⁵) about the relation of the mental to the physical that is stated and defended on the basis of the intentionality criterion.

In *Perceiving: A Philosophical Study* Chisholm states three conditions each of which is claimed to be a sufficient condition for a ‘simple’ or ‘non-compound’ sentence’s being an *intentional sentence*. There is no need for going into the details of the three conditions, but the ideas underlying them are by now familiar; first, certain *prima facie* psychological sentences (e.g., ‘Ponce de Leon sought the fountain of youth’) contain referring expressions whose success of reference is immaterial to the truth or falsity of the containing sentences; second, some putative psychological sentences (e.g., ‘I hope it will snow tomorrow’) have nested sentential clauses whose truth or falsity is independent of the truth-value of the containing sentences; and, finally, substitutivity of identity fails in sentences containing ‘verbs of propositional attitudes’. Now, according to Chisholm, intentionality thus defined is not exclusively a characteristic of psychological sentences, i.e., sentences ‘describing’ psychological phenomena; nor does he explicitly claim that all sentences about psychological phenomena are characterized by intentionality. His thesis rather is that any intentional sentence which *prima facie* is not about a psychological phenomenon can be ‘transformed’ into a non-intentional one. To quote Chisholm:

Let us say (1) that we do not need to use intentional sentences when we describe non-psychological phenomena; we can express all of our beliefs about what is merely ‘physical’ in sentences which are not intentional. But (2) when we wish to describe perceiving, assuming, believing, knowing, wanting, hoping, and other such attitudes, then either (a) we must use sentences which are intentional or (b) we must use terms we do not need to use when we describe nonpsychological phenomena.⁶

Ignoring (2-b), which is intended to prohibit the introduction of artificially defined expressions such as ‘deer-inclusive’ as in ‘His perceptual environment is deer-inclusive’ as a way of avoiding the intentional sentence ‘He perceives a deer’, the thesis – we shall call it ‘the Brentano-Chisholm thesis’ – comes to this:

The Brentano-Chisholm thesis: (1) Every physical phenomenon is describable by a nonintentional sentence. (2) No mental phenomenon involving psychological attitudes such as wanting, believing, and knowing can be described by a nonintentional sentence.

As stated, a full criterion of the mental or the physical cannot be extracted from this thesis, since it leaves open the possibility of mental phenomena that are describable nonintentionally. But an anti-materialist consequence follows from it nonetheless; the Brentano-Chisholm thesis implies that mental phenomena involving beliefs, wants, etc. are not physical phenomena.

In his entry on 'Intentionality' in *The Encyclopedia of Philosophy*, Chisholm makes things more explicit:

Some now believe that the thesis can be defended by reference to the language we use in describing psychological phenomena – that the sentences we must use in describing psychological phenomena have certain logical properties that are not shared by any of the sentences we must use in describing nonpsychological phenomena, and that these properties are correctly called intentional. If this view is true, then the basic thesis of physicalism and the unity of science is false.⁷

The thesis has now been extended to all psychological phenomena, making possible a full-fledged criterion of 'psychological sentence', and the anti-materialist implication is explicitly drawn. Clearly, if the situation is as described, the identity theory cannot be true, since physical events possess, and mental events lack, the property of describability by non-intentional sentences.

The key question in evaluating the Brentano-Chisholm thesis centers on the notion of 'description'; Chisholm's arguments in its support will be seen to involve a questionable assumption about this notion which would seem to weaken them considerably. First of all, how does he argue that every physical phenomenon is describable in nonintentional terms? Chisholm's procedure is to choose some typical intentional sentences which, intuitively, describe physical phenomena, for example, 'This weapon, suitably placed, is capable of causing the destruction of Boston,' and show that 'we can readily transform them into others'⁸ that are not intentional. But what does 'transform' mean? What counts as *correct* transformation and what counts as *incorrect* transformation? More specifically, if a sentence *S* is transformed into *S'* in the intended sense, what property of *S* must be preserved and transmitted to *S'*? There is a short and uninformative answer: *S* and *S'* must describe the same phenomenon. But this does not fully enlighten us; we must go on to ask: Under what conditions do two sentences describe the same phenomenon?

When we examine the actual cases of such ‘transformation’ in Chisholm, it becomes fairly clear that the required relationship between the two sentences *S* and *S'* is something like logical equivalence or synonymy; one must be ‘analyzable’ into the other, and their biconditional must be analytic. This interpretation gains strength when Chisholm’s argument for the other half of the thesis, i.e., that no psychological attitudes can be described nonintentionally, is examined. What Chisholm does here is essentially to show that no ‘definitions’ or ‘translations’ (Chisholm’s terms) untainted by intentionality are possible for such psychological sentences as ‘A man perceives an object *x*’, ‘*S* expects *E* to occur’, and sentences about believing.⁹ Thus, what is shown is *not* that no psychological phenomenon is describable nonintentionally, but the different thesis that no intentional psychological sentence has a nonintentional logical equivalent. The latter thesis by itself does not entail the former, and without the former the anti-materialist implication does not follow. And the entailment plausibly obtains *only if* one introduces the further premiss, tacitly assumed by Chisholm, that if two sentences are not logically equivalent they do not describe the same phenomenon, or, what is the same, sentences describing the same phenomenon are logically equivalent.

The problem of sentential descriptions of events – and the broader problem of the individuation of events – is one that has only recently come to philosophical attention and about which relatively little can be said that is uncontroversial.¹⁰ But it will be granted that logical equivalence is much too strong a requirement on the sameness of the event described by sentences; few will hesitate to say that ‘Socrates drank hemlock’ and ‘Xantippe’s husband drank hemlock’ describe one and the same event – or ‘phenomenon’ if you like. And the contemporary defenders of materialism have been careful to set their position apart from the already discredited claim that sentences about psychological events are definable in terms of, or translatable into, purely physicalistic expressions; their criterion of event individuation clearly excludes logical equivalence as a necessary condition of the identity of the event described. Thus, Chisholm’s argument at best shows the inadequacy of the thesis of physicalistic definability or translatability of mentalistic terms – that is, the so-called logical behaviorism – which was defended by some logical positivists in the 1930’s¹¹; it does not affect the more interesting form of

materialism, namely the identity theory, which asserts only a contingent identity of the mental with the physical; nor does it touch the sort of reductive physicalism and unity of science enunciated by Oppenheim and Putnam.¹²

Thus, the implementation of the first part of the Brentano-Chisholm program, i.e., that of finding a 'formal' or 'logical' property of psychological sentences, has no adverse effect on the identity theory, unless it is supplemented by an unrealistically stringent criterion of event individuation. Our point, in fact, seems applicable to all linguistic approaches to the mental-physical distinction: no reasonable definitions of 'psychological sentence' and 'physical sentence' can a priori exclude the identity theory in the form of the thesis that what is described by a psychological sentence is also describable by a physical sentence, unless it is accompanied by some unrealistic criterion of event individuation.

In any case, it is not certain that the program of defining 'psychological sentence' in terms of some logico-grammatical properties of sentences has succeeded or will succeed. On the face of it, if there were such properties, it would be a fantastic fact that would require a philosophical and linguistic explanation; we are used to thinking of the relation between language and the world as conventional, at least to an important extent, and it would be surprising if we discovered that there was a strict correspondence between some logical property of sentences and whether these sentences are *about* (*designate, describe, pertain to, etc.*) psychological phenomena. The mass of counterexamples¹³ against Chisholm's ingenious formulations of the intentionality criterion and the resulting complexity of Chisholm's more recent efforts¹⁴ reinforce this impression. Another important reason for the scepticism is the inherent vagueness in the notion of a sentence being *about* some psychological phenomenon (for example, is 'There are no pains' *about* some psychological phenomenon?), and unless more is said to clarify this concept it would be premature to expect an exact correspondence between it and some very precisely defined formal logical property of sentences.

III. THE CRITERION OF NONSPATIALITY

Apparently, the idea that what is mental lacks spatial dimensions originates with Descartes who, in a famous passage, wrote:

But although any one attribute is sufficient to give us a knowledge of substance, there is always one principal property of substance which constitutes its nature and essence, and on which all the others depend. Thus extension in length, breadth and depth, constitutes the nature of corporeal substance; and thought constitutes the nature of thinking substance. For all else that may be attributed to body presupposes extension, and is but a mode of this extended thing; as everything that we find in mind is but so many diverse forms of thinking.¹⁵

As far as I know, Descartes never explicitly said that minds are not *in* space, although certain doctrines of his, notably one that denies motion to mental substance, seem to entail it; in any event, all that the above passage implies is that minds do not have extension, i.e., three-dimensional bulk. As is well known, he said that “the seat of the soul” is in the pineal gland¹⁶, locating here “the actions of the mind.” It is interesting to note that Leibniz interpreted Descartes as saying that the mind has a spatial location in the body. In *New Essays on Human Understanding* Leibniz distinguishes three kinds of “ubeity” or “modes of existing somewhere”. The first kind is called ‘circumscriptive’ and characterizes the manner in which material bodies are in space; they are there “*punctatim*, in such wise that they are measures according to which we can assign the points of the thing placed corresponding to the points of space.”¹⁷ Of the second kind of ubeity and the location of the soul, Leibniz writes:

The second [kind of ubeity] is the *definitive*, when we can define, i.e., determine, that the situated thing is in such a space, without being able to assign the precise points or the peculiar places exclusive of what is there. Thus it has been considered that the soul is in the body, not supposing it possible to assign a precise point at which the soul or some portion of the soul is, without its being also at some other point.... It is true that Descartes desired to place narrower limits to the soul by locating it properly in the pineal gland. Nevertheless he did not dare to say that it is exclusively at a certain point in this gland; and this not being so he gains nothing, and it is in this respect precisely as if he gave it the entire body as its prison or place.¹⁸

The third kind of ubeity, finally, is the *repletive* and is attributed to God; but this does not concern us here.

Leibniz’ discussion does not make it clear whether the impossibility of giving the soul a precise location in the body indicates an epistemic fact about our ability to do so or an ontological fact about the objective indeterminacy of the soul’s spatial location. It is difficult to understand how any *object* can have an objectively indeterminate location within a region of space in which it is supposed to be located; if the soul is in the

body, it must be in either the lower half of the body or the upper half – or else it could not be in the body at all; if it is in the upper half, then it must be in either *its* upper half or lower half; and so on ad infinitum. On the other hand, if the definitive ubeyty of the soul is supposed to reflect our inability to find the exact location of the soul, it is unclear how we are able to locate the soul in the body to begin with.

Minds as substantive entities have largely disappeared from contemporary discussions of the mind-body problem, and those who concern themselves with the question of the spatiality of the mental attend chiefly to mental *events, states, processes*, and the like. Some have argued that mental events are not in space at all; for example, Norman Malcolm¹⁹ has claimed that it makes no sense to speak of “the bodily location of a thought” since there is, and can be, no intelligible test or criterion to determine the location of mental events in the body over and beyond a test for determining the location of the correlated neurological events; in saying this, Malcolm seems to have gone further than the classical dualists like Descartes and Leibniz. There are others who, like Leibniz, have argued that although mental events are located where one’s body is located, they cannot be given precise specific locations *within* the body.²⁰ On either view the identity theory would be impossible – that is, on the assumption that the appropriate bodily events with which mental events are to be identified have precise location in the body; under the identity theory, mental events too must have “circumscriptive ubeyty” if physical events do.

A third possible view is that mental events have a precise location in the body but at the present time we are unable to pinpoint their exact location; thus, on this view, mental events have “circumscriptive ubeyty”, objectively speaking, but epistemologically we can attribute to them only ‘definitive ubeyty’. Such a view is held by Anthony Quinton²¹ who thinks that mental events “have at least a rough position in space of a kind which here, as in other cases, scientific inquiry may render more precise.”²² But why should we think that mental events have even ‘rough position’ in space? Quinton’s answer is that they must have spatial location if they are properly to be ‘individuated’. Consider the following situation: Two persons *A* and *B* simultaneously experience qualitatively indistinguishable feelings of annoyance at a whistling noise in their vicinity. How do we justify, Quinton asks, our belief that there are here two experiences and

not one? Can we individuate the two mental events by the descriptions 'the experience caused by the noise in *A*'s hearing' and 'the experience caused by the noise in *B*'s hearing'? Quinton thinks not: "These two descriptions do not entail duality of reference. One and the same window can be broken by two bricks hitting it simultaneously."²³ Quinton's answer is that the two experiences are different in virtue of the different places at which they occur.

This argument needs examination. For Quinton's own solution to be adequate the descriptions 'the event *E* occurring at place *p*' and 'the event *E*' occurring at place *q*', together with the further assumption that *p* is different from *q*, must entail that *E* is different from *E*'; that is to say, it must be the case that events occurring at different places are different. But is this true? A house can be in both North Dakota and South Dakota; if this house goes up in fire, the fire occurs in both North Dakota and South Dakota. One might protest that the fire that burns *only* in one of the states is different from the fire that burns *only* in the other. This is true, but not relevant; it amounts to saying that 'the fire in this house that burns in North Dakota *but not in South Dakota*' and 'the fire in this house that burns in South Dakota *but not in North Dakota*' entail duality of reference. Thus, Quinton's spatially individuating descriptions must be strengthened thus: 'the experience occurring at place *p but not at q*' and 'the experience occurring at *q but not at p*' (or 'the experience occurring at *p but at no other place*', etc.). But this is a move of which we could have availed ourselves earlier without spatializing mental events; for, clearly, the two descriptions 'the experience caused by the noise in *A*'s hearing *but not in B*'s' and 'the experience caused by the noise in *B*'s hearing *but not in A*'s' entail duality of reference. Apart from this crucial weakness in Quinton's argument, the fact that mental events must be individuated with respect to spatially located objects does not necessarily imply that mental events themselves must be entities with spatial location; whatever 'individuation' might precisely mean, it would seem that we could individuate some properties with respect to material bodies exemplifying them ('the color of gold', 'the temperature of the sun's interior', and so on) without saying that these properties have location in space.

In any event, given that mental events are in space, why should they be located in the bodies of the persons to whom they occur? Why not say that all my mental events, and all yours, too, occur, say, at the tip of the

Washington Monument? Quinton's implied answer to this is that "the fact that their proximate causes and effects are in these bodies makes this [i.e. locating them in the bodies of their owners] the obvious thing to do"²⁴ and "[by] locating the mental causes of speech and behaviour in the body we at least locate both cause and effect somewhere in space and define a region within which a possible intermediary mechanism could be found."²⁵ This is an interesting idea, and I believe that causal considerations of this sort were operative in Descartes' thinking when he placed 'the actions of the soul' in the pineal gland.²⁶ The idea is that cause and effect must be spatially contiguous, or at least must have a spatially contiguous intermediary mechanism. But is this not an illegitimate extension to mental events of a condition properly applicable only to physical events? If mental events are not in space at all, the concept of spatial contiguity is not an intelligible notion for mental events; on the other hand, if they are to have location in space at all, the condition of contiguity could perhaps be invoked to locate them in the bodies of their owners²⁷. It is uncertain how compelling these causal considerations are, and, in any case, Quinton's argument fails to establish the assumption that mental events must have some location or other in space.

I believe that mental events do occur in space; I reach this conclusion not from causal considerations of the sort discussed above, although they certainly provide some motivation for it, but from more general considerations about the nature of events. It is convenient to think of an event as the exemplifying of a certain empirical property by an object at a time; a physical event is the exemplifying of a physical property by an object at a time; and a mental event, similarly, is the exemplifying of a mental property by an object at a time. All the objects we know that exemplify mental properties are biological organisms of certain complexity, especially human bodies. The location of an event is most naturally and conveniently defined as identical with the location of the object involved in it; that is, the event of an object x 's exemplifying a property P at time t is located where x is located at time t . From the assumption that mental events consist in human bodies' exemplifying mental properties it now follows that mental events occur where human bodies exemplifying them are located. Thus, the dualism of substance is frankly given up at the outset; to countenance nonphysical 'mental substances' as objects exemplifying mental properties and worry about

their spatial properties is already to have abandoned hope for a materialist account of man; under the monism of substance and the account of events and their locations given here, mental events turn out to have locations in the same sense in which physical events do. The location of an event, whether mental or physical, is derivative on the location of material objects.²⁸

What now of the argument that brain events with which mental events are identified can be given specific locations *inside* the body whereas on the present account the location of all mental events is the entire body? Even if mental events are conceded to have location in space, the identity theory still would be false since their location does not coincide with the location of brain events. The proper reply to this objection is that it all depends on what one means by 'brain event'. According to the present account, my body's being in pain is identical with my body's being in brain state *B*, and both these events are where my body is. The objection is based on the tacit assumption that my body's being in brain state *B* is identical with my brain's being in state *B* and that, as a result, these events are in my brain. But this we must deny; my body's being in brain state *B* is not the same event as my brain's being in state *B*. The event *x*'s exemplifying property *P* at *t* and *y*'s exemplifying property *Q* at *t'* are one and the same just in case *x* and *y* are the same object, *P* and *Q* are the same property, and *t* and *t'* are the same time. To be sure, it is necessarily true that if the event of my body's being in brain state *B* occurs, then so does the event of my brain's being in state *B*; but this does not show them to be the same event any more than the fact that necessarily if Xantippe's husband exists then Xantippe exists shows Socrates and Xantippe are one and the same person. My body's being in brain state *B* is a 'global event' of my body; it can no more be localized in some specific part of my body than the event of my body's having such-and-such average density. In fact, my body can have such-and-such average density without any specific proper part of my body having that particular density.

This account of events opens a way of formulating the Cartesian idea of spatiality into criteria of the mental and the physical. Given our analysis of events, what must be defined are the concepts of mental property and physical property; these will yield the concepts of mental event and physical event. As quoted earlier, Descartes wrote "All else

that may be attributed to body presupposes extension, and is but a mode of this extended thing". The following definition of 'physical property' would be in line with Descartes' idea.

Property *P* is a *physical property* if and only if that an object has *P* presupposes that it has extension.

The exact import of this criterion depends on how the controversial concept of presupposition is understood; it seems to me that implication or entailment will do here in place of presupposition. Also, 'has extension' may be weakened into 'has location in space'. Another refinement is necessary if being non-black and other such 'negative properties' are to count as physical properties; the needed modification can be accomplished by replacing the definiens by 'either that an object has *P* or that an object does not have *P* presupposes that it has extension'. Those who find the definiens of the above obscure on account of its use of the relation of presupposition (or implication or entailment) between 'that'-clauses could formulate it with respect to predicates:

Predicate *P* is a *physical predicate* if and only if '*x* is *P*' presupposes '*x* has extension'.

The possible changes and refinements noted earlier apply here also.

Could we formulate analogous definitions of 'mental property' and 'mental predicate'? Any object exemplifying a mental property is, as a matter of fact, a material body; but it is not necessarily a material body; on the other hand, an object exemplifying a mental property must be a concrete entity in time. Thus, one might try:

Property *M* is a *mental property* if and only if that an object has *M* presupposes that it is located in time but not that it has extension in space.

This lets in too much: *lasting for two hours* and *coming into being before the birth of Christ* turn out to be mental properties on this criterion. It is uncertain whether these and other possible counterexamples can be eliminated by adding further plausible conditions; it may well be that, within the limits set by the requirements of the identity theory, a workable criterion of the physical is the best we can hope for from considerations of spatiality.

Is the identity theory possible under criteria of the mental and the physical along these lines? Can mental properties be identical with

physical properties, given these criteria? The answer seems to be in the affirmative: two predicates P and Q can be such that ' x is P ' implies a sentence S , ' x is Q ' does not imply it, and yet P and Q designate the same property. For example, ' x is blue' implies ' x is not red'; ' x has the color of the sky' does not imply it; but 'is blue' and 'has the color of the sky' designate the same property. So the criteria by themselves do not preclude the identification of mental with physical properties, and the identity theory remains as a possible theory.

IV. THE EPISTEMIC CRITERION

As the quotations from Smart and Place at the outset of this paper indicate, most philosophers now seem to favor some sort of epistemic criterion; the concept of mental seems today more closely associated with such notions as 'privileged access', 'direct awareness', and 'privacy' than with the more abstract notion of intentionality and the basically unintuitive notion of nonspatiality. And the focus of philosophical discussion, at least as it concerns the question of materialism, too seems to have shifted from the rational and cognitive aspects of the mental to the phenomenal and affective aspects. Whereas, for Plato, rational thoughts were the paradigmatic activities of the soul, such lowly events as itches, toothaches, and twinges are now among those typically mentioned in philosophical discussion as mental events. And the identity theory is most often formulated as a thesis about sensations, or 'phenomenal events', to the effect that all sensations are physical processes in the brain.²⁹ In spite of the popularity of the epistemic criterion, however, one finds in the literature few overt attempts at a precise formulation of the criterion. In what follows, we shall use a proposal due to Heidelberger³⁰ as the vehicle of our discussion.

Stripped of inessential details and somewhat simplified, Heidelberger's criterion comes to this:

A singular statement that p is psychological if and only if there exists one and only one person S such that S may know that p even if S has no evidence for p or S may know that not- p even if S has no evidence for not- p .

The epistemic concept used here is that of *knowing without evidence*; the idea is that one's knowledge of one's own psychological states, at least

those sometimes called 'phenomenal states', is *direct* in that it is not based on, or inferred from, evidence or observation. The 'private' or 'privileged' character of such knowledge is captured in the qualification that there exists only one person who enjoys that sort of direct epistemic access.³¹ Heidelberger's criterion in effect says that a psychological statement is such that either it or its denial has these two properties. The reason for including reference to the denial of a statement is that some psychological statements such as 'Jones knows that the sky is blue', contain a nonpsychological clause, 'The sky is blue', which it entails but to which no one has direct private access. But the negation of such a statement, 'Jones does not know that the sky is blue', can be known directly and with privacy since 'Jones does not believe that the sky is blue', which implies it, can be so known.

It is well that Heidelberger restricts his criterion to singular statements, since no one person has direct and private access *vis-à-vis* generalized psychological statements such as 'Someone is in pain' and 'Exactly five persons are in pain'. But not even all singular psychological statements satisfy Heidelberger's criterion; as an example, consider ' $(A \ \& \ B) \vee (C \ \& \ D)$ ', where *A* is 'Abe is in pain', *B* is 'Bob is in pain', *C* is 'Charlie is in pain', and *D* is 'Dick is in pain'. Heidelberger's criterion is a good deal more plausible when applied only to 'simple' or 'atomic' singular statements; it should be clear that the epistemic property of private direct access cannot be used to give a direct definition of 'psychological' for compound and generalized statements.

But it is not at all certain that Heidelberger's criterion yields correct results even for noncompound singular statements which, intuitively, attribute some mental state to a person. A statement involving a non-referential singular term like 'Mr. Pickwick is in pain' is an obvious but unimportant counterexample; a more important kind of counterexample is a statement like 'The tallest man in this room is having a toothache'. This shows that Heidelberger's criterion does not satisfy what seems like a minimal condition on such a criterion that if ' $a=b$ ' is true, ' a is *F*' and ' b is *F*' are together psychological or together nonpsychological. Moreover, it is by no means clear that 'Jones is in pain' is psychological under the criterion; for just as Jones has to know that he is the tallest man in the room to know that the tallest man in the room is in pain, he must know that *he is Jones* to know that Jones is in pain.

The point being belabored is better appreciated if Heidelberger's criterion is recast in a formally correct form, since his formulation is formally defective; as stated, it is not a definition, as can be seen from the fact that it does not permit the elimination of the predicate 'is a psychological statement' from all contexts; for example, it does not permit the elimination of this term from 'The first statement in Descartes' cogito-argument is a psychological statement'. The following remedies this defect:

A singular sentence P is a psychological sentence if and only if there exists one and only one person S such that S may know that P is true without evidence or S may know that not- P is true without evidence.

In addition to the difficulties already mentioned which stand out more vividly under this formulation, there is another point to be noted: there is no guarantee here that logically equivalent sentences will be alike in point of being psychological or nonpsychological. This and some of the earlier difficulties are due to the special logical properties of contexts governed by epistemic terms; in view of the well-known problems and puzzles besetting such contexts, these difficulties are by no means unexpected. Not only do such contexts not admit of substitution of coreferential expressions but they do not allow interchange of logical equivalents; they are both opaque and ' L -opaque'. And in this respect they are worse behaved than contexts governed by the alethic modalities of necessity and possibility.

Consider the problem of using these opaque constructions in framing definitions; for example, consider the use of modally qualified terms in defining a class:

$$M =_{\text{def.}} \{x \mid \text{Necessarily, } x = \text{the morning star}\}.$$

This is a defective definition and does not succeed in characterizing a determinate class of entities; for the evening star is not a member of M , although the morning star, that is, the evening star, is a member of M . A similar problem arises for the following definition making use of a psychological intensional expression:

$$B =_{\text{def.}} \{x \mid \text{Jones believes that } x \text{ has two daughters}\},$$

where we assume Jones to be a person who believes that President Nixon has two daughters but does not believe that the vice president under President Eisenhower has two daughters. And we should expect the same difficulty with the definition of 'mental' that makes use of the opaque construction '*S* may know without evidence that *P* is true'; given two names or descriptions of a sentence, the sentence may be psychological or nonpsychological under Heidelberger's criterion depending on which name or description is used.

The opacity of constructions involving epistemic terms constitutes a basic difficulty for any epistemic definition of 'mental'; we shall later indicate a way of alleviating this difficulty, but it is surprising that no notice has been taken of this problem by philosophers who have favored an epistemic criterion of the mental. This is not to say that philosophers have been unaware of the opacity of epistemic constructions in connection with the mind-body issue; in fact, this opacity has been used by the supporters of the identity theory to answer a familiar type of objections. These objections, which are conveniently called 'epistemic objections', are based on the observation that there are certain epistemic properties, such as privacy and direct accessibility, which characterize mental events – at least some of them – but not any physical event. Thus, one familiar line goes like this: "I am directly aware of my pains but not of my brain events; therefore, pains are not brain events." The reply to this is equally familiar: "The context 'I am directly aware that...' or 'I am directly aware of...' is opaque and does not permit substitution of coreferential expressions."³² But the identity theorist whose conception of the mental is essentially epistemic cannot avail himself of this line of defense without inviting the further challenge that if 'directly aware' is indeed opaque, his concept of mental event is not a determinate concept and as a result the identity thesis itself fails to make a determinate assertion.

Let us now try to formulate a workable epistemic criterion of the mental. Our formulation will be ontological rather than linguistic, and will give us a criterion of 'phenomenal property'. Roughly speaking, a phenomenal property is a property associated with a mental event in virtue of which the event is 'directly' apprehended; mental events clearly have properties, such as causal properties, that are not directly apprehended. A phenomenal event is an event which consists in the exemplifica-

tion of a phenomenal property at a time; the notion of 'phenomenal event' is intended to cover events like sensations, images, and occurrent thoughts, but not those psychological states involving dispositions like desires and motives; it will be granted that our epistemic position *vis-à-vis* these states is not in principle superior to anyone else's.

Property *F* is a *phenomenal property* if and only if for any *S* if *S* exemplifies *F*, *S* is the only person who is directly aware that *S* exemplifies *F*.

The variable '*F*' appears both inside and outside the opaque construction involving 'directly aware'. It might be argued that while the definiens is satisfied when 'pain' replaces '*F*' it is not satisfied when '*F*' is replaced by 'the mental property nomically coextensive with the stimulation of the hypothalamus' even if this latter expression refers to pain. Similar comments apply also to the variable '*S*'. To minimize this sort of difficulty we can use a simple device analogous to one used by E. J. Lemmon.³³ Recall the defective definition of class *M* above; this can be amended thus:

$$M =_{\text{def.}} \{x \mid (\exists y) (x = y \ \& \ \text{necessarily } y = \text{the morning star})\}.$$

Using a similar device we have:

Property *F* is a *phenomenal property* if and only if for any *S*, if *S* exemplifies *F*, there exists *S'* and *F'* such that *S* = *S'*, *F* = *F'*, and *S* is the only person who is directly aware that *S'* exemplifies *F'*.

It is not certain that this device eliminates all the objectionable features of the opaque construction in question, but it helps us avoid the obvious pitfalls.

I have used 'directly aware' without a detailed explanation of its meaning; in fact, I am using it as a catchall term for privileged epistemic access. My formulation of the epistemic criterion, therefore, is best taken as giving only a form in which specific epistemic criteria can be stated; it will yield different criteria depending on exactly how 'directly aware' is understood. In particular, the clause '*S* is directly aware that *S'* exemplifies *F'*' can be replaced by others to suit one's predilections; for example, one could use the notion of 'knowledge without evidence' by replacing '*S* is directly aware' with '*S* is in a position to know without evidence'; or

we could use the weaker condition 'if *S* believes that *S*' exemplifies *F*', then *S* would be maximally justified in this belief'.³⁴ There are two features of my formulation which may be worth noting: first, the use of a device to circumvent the opacity of epistemic expressions, and second, the fact that the criterion is stated for properties, not for such more inclusive units as states and propositions.

How this concept of phenomenal property can be extended into a general concept of mental property is a question into which we cannot enter here; but a hint as to the direction such an extension might take is contained in the idea that the ascription of a mental property to an object not directly involving a phenomenal state implies that the object to which it is ascribed has been or is or will be in some phenomenal state if certain specified conditions are realized; thus, to say that a person *wants* some object *x* is to say, in part, that he will be *pleased* when he believes that he has *x*.³⁵ In fact, the property of being directly aware used in our formulation of the epistemic criterion does not seem to qualify as a phenomenal property under it; it usually is not the case that when I am directly aware of something, I am directly aware that I am directly aware; there is a well-known infinite regress that threatens to get started here. Thus, direct awareness (whatever its precise interpretation may be) is likely to be counted in the broader class of mental properties, not in that of phenomenal properties.

V. CONCLUDING REMARKS

The problem of characterizing the intentionality of sentences and its possible connection with the psychological subject matter of these sentences is a fascinating problem; but as I noted earlier, there is reason to be sceptical about the success of this project, and it would be wrong to pin on it our hope for a criterion of the mental. I think we can be rather more hopeful with regard to the other two major criteria. The epistemic criterion could give us a viable concept of the mental – at least, the concept of the phenomenal which is sufficient for stating the currently fashionable versions of the identity theory; and the spatiality criterion could give us a workable concept of the physical. It seems to me that this is the correct approach; the epistemic criterion, which is intended to capture one facet of the concept of consciousness, seems well suited for

delineating those events which the identity theorist is anxious to identify with brain events; and the spatiality criterion of the physical is consonant with our idea of the physical as that which belongs in the intersubjective space-time network.³⁶

I shall conclude with a few remarks about the two approaches toward a criterion of the mental that we have distinguished, namely the ontological and the linguistic. We should note first that an ontological criterion does not automatically yield a linguistic criterion, nor conversely. That is, we cannot simply assume that an expression is mental if it describes, refers to, or designates something that is mental, or, conversely, that if something is designated or described by a mental expression it must be a mental entity. The very possibility of the identity theory is contingent on the falsity of the first; or consider 'the nonphysical event nomically coextensive with the stimulation of the hypothalamus'. Similar constructions will show that the second assumption is equally faulty. In fact, the idea that a psychological sentence 'describes' a psychological phenomenon is not a clear idea at all. Does the sentence 'There are no pains' describe a psychological phenomenon? I think it is a mistake for someone looking for a linguistic formulation to attempt to tackle *sentences* as the basic units of classification into mental and nonmental. We have two conflicting intuitions about what is to count as a psychological sentence which are not easily reconciled: on the one hand, a psychological sentence is one that contains psychological expressions essentially, and on the other, a psychological sentence is one that, in some sense, describes a psychological phenomenon. 'There are no pains' will count as psychological under the first, but presumably not under the second; and if the identity theory is true, sentences that are not psychological in the first sense will describe psychological phenomena. I think we should be better off tackling *predicates*, rather than sentences, in formulating a linguistic criterion of the mental or the physical. As will be recalled, all the authors we have mentioned in this paper who take the linguistic approach attempt to work with full sentences.

Another pitfall with the linguistic approach is the temptation to use some sort of logical relationship between sentences as a closure condition on the class of psychological sentences as a way of enlarging it; thus, Chisholm explicitly accepts at one point the principle that a consistent sentence implying a psychological sentence is itself psychological.³⁷ This

would be a tempting course if, as is likely, a given definition of 'psychological sentence' does not directly classify, e.g., 'Jones knows that snow is white' as psychological while it does so classify 'Jones believes that snow is white'. There are other logical closure conditions that at the first blush look equally tempting. But we must be extremely cautious with these conditions; even Chisholm's seemingly plausible closure condition yields what seems to be an unacceptable consequence that 'There are no men' is psychological since it implies 'No man believes that unicorns exist' which is psychological by Chisholm's intentionality criterion. In another place Chisholm accepts the condition that a compound sentence is psychological if one of its components is.³⁸ The combination of these two conditions would lead to the disastrous consequence that every sentence is psychological if at least one is.

I point out these difficulties with the linguistic formulations because they seem to indicate that the notion of 'psychological sentence', even if we had a clear notion of 'psychological predicate', is neither very clear nor very useful. And in view of the difficulty alluded to earlier of explicating the notion of a sentence describing an event or phenomenon, much more would have to be said to clarify the linguistic versions of materialism which typically assert something like "Everything described by a mentalistic sentence can be described by a physicalistic sentence". An ontological criterion that directly deals with mental and physical entities is free from these difficulties, and should therefore be preferred to a linguistic one, *ceteris paribus*.³⁹

The University of Michigan

NOTES

¹ *Philosophy and Scientific Realism*, London 1963, p. 89.

² 'Is Consciousness A Brain Process?', reprinted in John O'Connor (ed.), *Modern Materialism: Readings on Mind-Body Identity*, New York 1969, p. 22.

³ Carl G. Hempel distinguishes between the 'linguistic' thesis and the 'ontological' thesis of physicalism in 'Reduction: Ontological and Linguistic Facets', in S. Morgenbesser *et al.* (eds.), *Philosophy, Science, and Method*, New York 1969.

⁴ Franz Brentano, 'The Distinction Between Mental and Physical Phenomena', from *Psychologie vom empirischen Standpunkt*, first published in 1874. The quotation is taken from Harold Morick (ed.), *Introduction to the Philosophy of Mind*, Glenview, Ill., 1970, pp. 119-120.

⁵ In his *Perceiving: A Philosophical Study*, Ithaca 1957, p. 172.

⁶ *Ibid.*, pp. 172-173.

- ⁷ Ed. Paul Edwards, New York 1967, Vol. 4, p. 203.
- ⁸ *Perceiving: A Philosophical Study*, p. 173.
- ⁹ See his 'Sentences About Believing', *Proceedings of the Aristotelian Society* 56 (1955–56) 125–48, as well as *Perceiving: A Philosophical Study*, *ibid.*
- ¹⁰ See, e.g., Donald Davidson, 'The Logical Form of Action Sentences', in N. Rescher (ed.), *The Logic of Action and Decision*, Pittsburgh 1967, and 'The Individuation of Events', in N. Rescher *et al.* (eds.), *Essays in Honor of Carl G. Hempel*, Dordrecht 1969. R. M. Martin, 'On Events and Event-Descriptions', in J. Margolis (ed.), *Fact and Existence*, Oxford 1969.
- Jaegwon Kim, 'Events and Their Descriptions: Some Considerations', in N. Rescher *et al.* (eds.), *op. cit.*
- ¹¹ See, e.g., Rudolf Carnap, 'Psychology in Physical Language', in A. J. Ayer (ed.), *Logical Positivism*, Glencoe, Ill., 1959. First published in German in *Erkenntnis* 3 (1932/1933).
- ¹² Paul Oppenheim and Hilary Putnam, 'Unity of Science as a Working Hypothesis', in H. Feigl *et al.* (eds.), *Minnesota Studies in the Philosophy of Science*, Vol. 2, Minneapolis 1958.
- ¹³ See, e.g., Herbert Heidelberger, 'On Characterizing the Psychological', *Philosophy and Phenomenological Research* 26 (1966) 529–36; W. Gregory Lycan, 'On Intentionality and the Psychological', *American Philosophical Quarterly* 6 (1969) 305–11.
- ¹⁴ For some of Chisholm's recent formulations, see his 'Brentano on Descriptive Psychology and the Intentional', in E. N. Lee and M. Mandelbaum (eds.), *Phenomenology and Existentialism*, Baltimore 1967; and 'On Some Psychological Concepts and the "Logic" of Intentionality', in Hector-Neri Castañeda (ed.), *Intentionality, Minds, and Perception*, Detroit 1967.
- ¹⁵ *Principles of Philosophy*, I, liii.
- ¹⁶ In his letter to Meissonier Descartes wrote: "Mon opinion est que cette gland [the pineal gland] est le principal siège de l'âme, et le lieu où se font toutes nos pensées." Victor Cousin, *Oeuvres de Descartes*, Paris, Vol. 8, p. 200.
- ¹⁷ Bk. ii, Ch. xxiii, Sec 21.
- ¹⁸ *Ibid.*
- ¹⁹ 'Scientific Materialism and the Identity Theory' reprinted in O'Connor, *op. cit.*
- ²⁰ Jerome A. Shaffer writes: "The most sensible answer would be that [the thought] occurs wherever you are when you have that thought. If you are in the library when you have that thought, then the thought occurs in the library. But it would be utterly unnatural to ask where inside your body the thought occurred...." *Philosophy of Mind*, Englewood Cliffs, N.J., 1968, p. 48.
- ²¹ In his comprehensive and illuminating discussion of the mind-body distinction 'Mind and Matter', in J. R. Smythies, ed., *Brain and Mind*, London 1965.
- ²² *Ibid.*, p. 214.
- ²³ *Ibid.*, p. 211.
- ²⁴ *Ibid.*, p. 213.
- ²⁵ *Ibid.*, p. 214.
- ²⁶ Descartes' stated reason for choosing this particular location is that it is the only organ in the brain that does not come in pairs and that this is required to explain the unity of mental experiences. See *Passions of the Soul*, Part I, Article xxxii.
- ²⁷ This consideration makes it difficult to understand why Leibniz, who abjured all causal connections between mind and body, should feel it necessary to have the soul in the body.

²⁸ For more details see my 'On the Psycho-Physical Identity Theory' and 'The Logic of the Identity Theory' (with Richard B. Brandt) reprinted in O'Connor, *op. cit.* I give a fuller account of events in 'Causation, Nomic Subsumption, and the Concept of Event', forthcoming.

²⁹ Smart's influential first paper on the identity theory is entitled 'Sensations and Brain Processes', reprinted in O'Connor, *op. cit.*

³⁰ *Op. cit.* Harold Morick states two formulations of the epistemic criterion in his 'Introductory Essay' in his *Introduction to the Philosophy of Mind, op. cit.*

³¹ For excellent discussion of the various senses of 'privileged access' see William P. Alston's 'Varieties of Privileged Access', forthcoming.

³² See, e.g., Thomas Nagel, 'Physicalism' reprinted in O'Connor, *op. cit.*, p. 105. A fuller discussion of epistemic objections to the identity theory is contained in Richard B. Brandt and Jaegwon Kim, 'The Logic of the Identity Theory' in O'Connor, *op. cit.*, pp. 234-36.

³³ In 'A Theory of Attributes Based on Modal Logic', *Acta Philosophica Fennica* 16 (1963) 95-122.

³⁴ This corresponds to Alston's condition of 'truth-sufficiency'. For other interesting variations, see his paper cited in Note 31.

³⁵ This as well as my formulation of the epistemic criterion (except for the refinement to avoid the difficulties of opaque constructions) were stated in Brandt and Kim, *op. cit.*

³⁶ This is the sense of 'physical₁' explained by P. E. Meehl and Wilfrid Sellars in 'The Concept of Emergence', H. Feigl *et al.* (eds.), *Minnesota Studies in the Philosophy of Science*, Vol. 1, Minneapolis 1956, p. 252.

³⁷ In 'Brentano on Descriptive Psychology and the Intentional' and in his 'Rejoinder' to R. C. Sleight's 'Comments', in Castañeda, *op. cit.* The condition is stated in terms of 'intentional', not 'psychological'.

³⁸ In *Perceiving: A Philosophical Study*, p. 172. Chisholm states this in terms of 'intentional', not 'psychological'.

³⁹ I am indebted to William P. Alston and Alvin I. Goldman for helpful comments.