

Introduction

The metaphysical theory of mind I advocate has three primary component theses:

- 1. Psychological statements are frequently true.
- 2. The only entities required for the truth of psychological discourse are persons and other sentient creatures together with sets built up out of such entities.
- 3. Collectively, the various generalizations of a psychological theory, be it folk or scientific, implicitly define the psychological predicates occurring within that theory.

There are also five subsidiary negative theses that figure in the development of my overall position:

- 4. There are no mental events.
- 5. Sensation is not relational: There are no sensory objects.
- 6. The propositional attitudes are not relational: There are no objects of the attitudes, be they abstract propositions or sentences in an inner or an outer language.
- 7. Cognitive psychology is not committed to the existence of mental representations.
- 8. Neither modal arguments nor the argument from knowing 'what it is like' establish that anything is left out by a physicalist approach to the subjective aspects of sensory experience.

Of the eight theses, the first is given the least defense. This is because I believe that most philosophers do not need to be convinced of its truth and also because I have little to add to standard criticisms of eliminative materialism, that is, the view that present-day psychological statements are infected with radical error.

The structure of my argument is basically as follows. It is widely held that there really are such entities as mental events – for example, beliefs and desires. At the beginning of Chapter 1 (which is written

1 This is to oversimplify a little; in certain cases space-time regions are also required.



	Psychophysical event dualism	Token identity theory	Eliminative materialism	Operator theory
Psychological statements are frequently true	Yes	Yes	No	Yes
There are mental events	Yes	Yes	No	No
There are nonphysical mental events	Yes	No	No	No

jointly with Terence Horgan) it is shown that this view generates a problem: that of finding a satisfactory niche for mental events in nature as a whole. Many philosophers solve this problem by endorsing the token identity theory (the theory that every mental event is strictly identical with some physicochemical event). However, in Chapter 1, it is argued that the token identity theory is false (as is the closely related materialist view that mental events are complex entities that are composed of physicochemical events but that are nonetheless distinct from the aggregates or fusions of their parts). What to do? There are only two courses of action left open: Become a psychophysical event dualist or deny that mental events really exist. In Chapter 2, I argue for the latter alternative. This smacks strongly of eliminative materialism, a view I reject. So the argument appears to lead nowhere. But there is a way out, I suggest, since the claim that mental events do not exist does not entail that eliminative materialism is true. Rather it entails only that either eliminative materialism is true or psychological statements have a misleading grammatical form and, contrary to appearances, do not require for their truth the existence of mental events (just as 'Paul died for the sake of his country' does not require for its truth that there really exist such an entity as a sake). In Chapter 3, I embrace the second of these disjuncts, and I present the outline of a metaphysical theory of mind that has the commitments expressed in thesis 2. For reasons that will become clear later, I call this theory the "operator theory." The table above summarizes how the operator theory resembles and differs from the usual three alternative theories on the issues of truth and mental events.

The operator theory repudiates not only mental events but also mental objects – for example, pains and images. These entities face



the same sort of problem as the one presented in Chapter 1 for mental events. They also encounter all sorts of well-known philosophical puzzles of their own. The underlying reason for these puzzles, according to the operator theory, is straightforward enough: Psychological statements putatively about mental objects are, in reality, free of any such commitment. There are no mental objects, just as there are no mental events. This point is given a preliminary defense in Chapter 3.

Chapters 4, 5, and 7 are concerned primarily with the detailed development of the operator theory in the context of the statements of folk psychology. I begin with statements "about" bodily sensations such as pain; I proceed next to statements "about" perceptual sensations; and I continue on in Chapter 7 to statements "about" belief, desire, and other propositional attitudes. I argue that all of these statements can be regimented in the same general way and that there is no need to invoke such entities as sense data, mental sentences, or abstract propositions in order to account for the truth of folk psychological discourse.

So far in my explanation of the operator theory I have said nothing about the meaning of psychological discourse and the status of sensory qualia. Nevertheless, as I indicated at the beginning, I have definite views on these matters. On the issue of the status of qualia, I offer, in Chapter 6, what I hope is a compelling defense of thesis 8. On the issue of meaning, I opt for a broadly functionalist approach to psychological predicates. My primary concern here is to show how a version of conceptual functionalism can be developed that goes hand in hand with my general semantical views and that also avoids any metaphysical commitment to events (including states and conditions) or even to properties. This seems to me worthwhile in part because conceptual functionalism sometimes seems to be an isolated position set apart from other views on the semantics of everyday nonpsychological language, and in part because it is not obvious what a satisfactory reconstruction of functionalism would look like (or even whether it is possible) within the metaphysical confines of my position. Parts of Chapters 4, 5, 7, and 8 address these topics.

In the concluding chapter, my discussion of the operator theory moves from folk psychology to cognitive psychology. It is widely held that cognitive psychology is committed to the existence of mental representations (be they quasi-linguistic, quasi-pictorial, or



otherwise). In opposition to this view, I present an interpretation of cognitive psychology that avoids any commitment to mental representations, and I maintain that from the point of view of metaphysics this interpretation is preferable to the standard one.

The view that emerges from my discussion is a form of physicalism with respect to the mental in that it denies that there are any nonphysical mental entities. But it is a physicalism that avoids both the radical extremism of eliminative materialism and the pitfalls of the token identity theory. Admittedly the truth conditions I state for psychological discourse involve sets, and this may not sit well with physicalists who are die-hard nominalists. Still, as Quine has stressed, sets do at least have the virtue of clear, uncontested identity conditions. Moreover, they seem needed in formal treatments of truth conditions generally. My hope, I should add, is that this commitment is merely apparent and that sets are ultimately eliminable both in the truth conditions for psychological discourse and elsewhere.² However, I make no attempt to argue for this view in the present book. Overall, it seems to me, the operator theory has great systematic unity and great simplicity. For these reasons, it is, I believe, the most promising metaphysical theory of mind available to the physicalist.

The metaphysical views presented in this book rest on an important background assumption. Space limitations prevent me from offering a full defense of this assumption here. But I think I ought to explain what the assumption is and, in general terms, why I make it.

With one qualification to be stated shortly, I assume that the existential quantifier is a constant indicator of reference that provides us with the means to determine the ontological commitments of what we say. This view of metaphysics is, of course, that of Quine.³ The reasoning behind it can be summarized as follows. Everyday English is full of expressions that appear to be singular terms but that are not used referringly, for example, 'the average family', 'Pegasus', 'the behalf of Amy', 'the equator'. So the or-

² For a promising account of how such an elimination might go, see Terence Horgan, "A Nominalistic Theory of Truth," *International Logic Review*, forthcoming.

³ See W. V. O. Quine, "On What There Is," in From a Logical Point of View (New York: Harper & Row, 1963); idem, Word and Object (Cambridge, Mass.: MIT Press, 1960).



> dinary assertion of an English sentence containing an apparent singular term does not commit its speaker to the existence of an entity corresponding to that term. How, then, are we to determine the ontic commitments of given utterances? Quine's position, in essence, is that we must translate the utterances into a formal language incorporating the existential quantifier ' $(\exists x)$ '. We then commit ourselves ontologically by asserting sentences that contain ' $(\exists x)$ ' or that logically imply sentences containing ' $(\exists x)$ '. In taking this view, Quine makes no attempt to explicate or analyze the ordinary term 'exist'. Quine's strategy is rather to introduce a symbolic expression, namely the existential quantifier, which formally captures the standard existentially committing use of this term without further illuminating its meaning. The test of a set of translations of a given body of discourse, θ , into the language of quantification is simple enough: Since the formalized sentences are intended to represent perspicuously the truth conditions of the informal sentences making up θ , any proposal that fails to preserve any of the various logical entailments and nonentailments that obtain between the members of θ must be rejected as unsatisfactory.

> The qualification I referred to earlier pertains to any existential quantifier having a component variable the substitution instances for which are linguistic items other than singular terms. I hold that such a quantifier is not an indicator of reference, and hence I also hold that its use need not be existentially committing. In the case of the quantifier ' $(\exists x)$ ', the substitution instances for the variable 'x' are usually taken to be genuine singular terms, that is, expressions that refer to entities in the world. It is in virtue of this fact that the use of ' $(\exists x)$ ' is taken to indicate ontic commitment. In the case of ' $(\exists F)$ ', however, in, say, ' $(\exists F)Fa$ ', the substitution instances for the variable 'F' are predicates, and predicates are obviously not singular terms. This point is discussed further in Chapter 3, where it is suggested that quantifiers such as ' $(\exists F)$ ' and '(G)' be read substitutionally.

It may be objected that once ' $(\exists F)$ ' is interpreted substitutionally there is nothing to prevent us from interpreting any quantifier we wish substitutionally, in which case the sometimes complex analyses of logical form presented in this book become otiose. However, the attempt to interpret ' $(\exists x)$ ' substitutionally is fraught with difficulty. One major problem is this: If we hold that ' $(\exists x)Fx$ ' is true just in case some instance of 'Fx' is true, then not only is it



difficult to grasp what the right side of the biconditional asserts if it does not involve objectual quantification over linguistic items, but there also remains the task of spelling out adequate truth conditions for atomic sentences of the form 'Fx'. This task, moreover, is decidedly nontrivial; for the truth conditions will have to show how the world makes the atomic sentences true, and they will have to do so without reintroducing an ontological commitment to the entities that the substitutional approach is trying to avoid. In particular, the truth conditions cannot take the substitution instances of 'x' in sentences of the form 'Fx' to be names so that 'Fa', say, is true if, and only if, some individual named by 'a' belongs to the satisfaction set of 'F'.

I do not claim that these difficulties cannot be overcome. But I do not myself see how to overcome them. And minimally I think they make it clear that a glib appeal to a substitutional interpretation of the quantifier ' $(\exists x)$ ' whenever there appears to be unwanted ontic commitment carries no weight.

In concluding this introduction, I want to show how the approach to metaphysics I have sketched undercuts two general criticisms of my views that have cropped up in conversation. I have been asked repeatedly what relevance my views have for cognitive psychology. Even if I repudiate events, it has been said, cognitive psychologists, under the influence of the computer model, certainly do posit information-bearing events in people and other cognitive systems. And why shouldn't they? Why, from their own point of view and for scientific purposes, should cognitive scientists and researchers in artificial intelligence engage in the logical maneuvers I defend in an attempt to avoid metaphysical commitment to information-bearing states and events?

This criticism rests on a serious misunderstanding of what I am doing. First, my views are metaphysical. They are intended to have absolutely no relevance to the day-to-day practices and assertions of cognitive scientists. I am happy to grant that such scientists have nothing at all to learn from what I say that could or should influence their methods or theories. To suppose otherwise would be just as absurd as supposing that an ordinary person should change his or

⁴ For more on the problems associated with a substitutional interpretation of all quantifiers, see Dale Gottlieb, *Ontological Economy* (New York: Oxford University Press, 1980), pp. 48–50.



> her day-to-day assertions about the mind in the light of my discussion. Second, the fact that cognitive psychologists propound theories that appear from their superficial grammatical form to be committed to the existence of representations, events, states, and so on does not entail that those theories really are so committed. That issue – the metaphysical issue – is not one for the cognitive scientist but rather for the philosopher. The scientist interested in the mind presents theories of how we perceive, how we remember, how we understand, and so on, which he or she believes to be true. The descriptive metaphysician, in my view, then, regiments these theories (in the manner I have indicated) in an attempt to discover what there must be in the world in order for them to be true. Third, the empirical evidence for the claim that cognitive scientists posit information-bearing events or states is really only evidence for the claim that cognitive scientists propound theories that putatively refer to or quantify over information-bearing events or states. Putative reference is not the same as genuine reference. Whether the cognitive theories must be interpreted as really referring to (or quantifying over) events and states is not an issue that can be settled by examining the intentions, assertions, or scientific procedures of cognitive scientists. Instead it is an issue for logical analysis.

> The second criticism I want to address is that the operator theory, as I elaborate it, is merely a linguistic dodge that reveals nothing about the mind in particular and is subject to no significant constraints. This criticism reveals another serious misunderstanding. The operator theory does not consist of purely syntactic, formal analyses; it has an important semantic dimension. In this respect, it is like any other metaphysical theory (given my conception of metaphysics). Moreover, the semantic aspect of the operator theory places well-defined constraints on its applicability. If there are logical entailments among psychological sentences that the theory cannot justify via the truth conditions it assigns to those sentences, then the theory is just plain wrong. If there are nonentailments among psychological sentences that the theory converts into entailments via its assigned truth conditions, then again the theory is wrong. As to whether the strategies the theory employs may be applied elsewhere, outside of either the mind or the realm of events, I regard that as an open question. It seems to me that the operator theory provides us with a well-motivated and cohesive view of the mind and that the strategies it rests on find their natural home in



the contexts I examine. But I do *not* wish to rule out the possibility that similar strategies may be developed in other areas. I merely note that any philosopher interested in applying an approach like mine to the task of avoiding further entities must *show* (not merely assert) that all the types of statement nonphilosophers make "about" those entities can be reconstructed without reference to or quantification over them – the test, as before, being whether there are inferences among the relevant statements that cannot be justified by reference to the assigned semantical analyses.



1

An objection to the token identity theory

with TERENCE HORGAN

Philosophers have generally agreed that there are concrete mental events, for example, my now thinking of England, your remembering your last meal, my feeling pain, your desiring a week in the sun. There has been considerable disagreement, however, about where to locate mental events in nature as a whole. The modern debate on this issue originates, of course, with René Descartes. According to Descartes, mental events take place in spiritual substances that are unextended, nonspatial, and immaterial. Minds, on Descartes's view, just are these spiritual substances, and though they are nonphysical, they are related in a special and close way to the physical bodies with which they are associated via a two-way causal interaction between mental events and bodily changes.

Descartes defended this view of the mind by appeal to famous arguments from doubt and possibility, arguments that need no repetition here. The view has been influential in philosophy mainly because of these arguments, but also to a lesser degree because historically many philosophers have shared the intuition that the mind and its contents are somehow intrinsically different from the rest of nature.

Today Cartesian dualism has diminishing support. Its decline can be traced to a number of factors, central among which have been a skepticism about the intelligibility of the concept of an immaterial substance and a growing belief, fostered by recent work in the physical sciences, that the physical world is causally closed. This last belief, if true, precludes any causal interaction between spirit and matter.

1 See, e.g., Meditations on First Philosophy, first published in 1641.



But if the physical world is causally closed, then where within it are mental events to be found? One straightforward and widely accepted answer is that every concrete mental event is strictly identical with some concrete physicochemical event. The view expressed in this answer, which has come to be called the "token identity theory," has no difficulty in explaining the special nature of the connection between a person's mind and a person's body; furthermore, it sidesteps worries about the intelligibility of the concept of a spiritual substance, and it has the immediate advantage of ontological simplicity. Nonetheless, we believe that it should be rejected. The purpose of this chapter is to present and defend our major objection.

I

In accordance with common recent practice, we shall use the rubric 'event' in a broad sense, to include not merely changes but also states, processes, and the like. Events, as we shall here construe them, are concrete entities, or tokens. If we mean to speak of the types of which events are tokens, we shall explicitly use the term 'event-type'. (Occasionally we shall use the term 'event/state' rather than 'event', in order to emphasize that the entities we are speaking of can have longer than momentary duration and that these entities can be — or have parts that are — relatively static during their duration.)

Now if mental events exist, then for any creature c who has mentality, there is a nonempty set M(c) containing all and only the mental events of which c, at one time or another during the creature's lifetime, is the subject; we shall call this c's mentality set. We shall take the contents of M(c) to include not only events of the kind that are apparently posited by common-sense psychology ("folk psychology"), but also mental events of any additional kinds that would be posited by an ideal theoretical psychology.

For any creature c with a nonempty mentality set M(c), we shall say that a set of events P(c) is a physical causal isomorph of M(c) (for short, a PCI) if, and only if, (1) every member of P(c) is a physicochemical event of which c is the subject, and (2) there is a one-to-one relation R between the events in P(c) and the events in M(c) such that (a) each event in P(c) is simultaneous with its R