INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9013892

Physicalism & psychology

Endicott, Ronald Patrick, Ph.D.

The University of Michigan, 1989

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

PHYSICALISM & PSYCHOLOGY

by

Ronald Patrick Endicott

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Philosophy)
In the University of Michigan
1989

Doctorial Committee:

Professor Jaegwon Kim, Co-Chairman
Assistant Professor William Taschek, Co-Chairman
Professor Edward E. Smith
Assistant Professor Stephen Yablo

Dedicated to the Memory of my Mother,

Who Instilled in me both the Appreciation of Nature,

and the Passion to Know our Place within it

ACKNOWLEDGMENTS

Initial thanks go to my family and friends, whose presence helped to make graduate life a worthwhile affair. I should also like to thank the many graduate students who provided a constant source of stimulation which, for better or worse, affected my overall intellectual development, chief among them Don Demetriades, Gary Ebbs, Don Loeb, David Reed-Maxfield, Joel Richeimer, and Michael Winstrom. Another word of thanks goes to George Mavrodes, who read the entire manuscript, and was helpful in many ways during my stay in Ann Arbor; and also to Robert Pachella, who more than any other showed a genuine concern for graduate student education. Finally, special thanks go to the members of my committee, Jaegwon Kim, William Taschek, Stephen Yablo, and Edward Smith, who made invaluable contributions and suggestions throughout this thesis.

PREFACE

My project can be seen as a natural extension of the "Cartesian program" advanced in many quarters (for example, linguistics, psychology, and artificial intelligence), that is, as an extension of the general approach which seeks to describe an autonomous domain of psychological entities and laws. Here the autonomy is preserved by supplementing the Cartesian program with a sophisticated Cartesian metaphysics. Thus, as we shall see in the course of this study, I resist any doctrine of physicalism which would challenge the metaphysical autonomy of psychological theory.

To help locate my view within the market place of recent philosophy, the position I advance has some affinity with what Paul Feyerabend calls "theoretical pluralism" [Feyerabend,1963b], only I accept a pluralism at the inter-theoretic level whereas Feyerabend restricts his pluralism to an intra-level proliferation of basic physical theories. Consequently, my view more closely resembles what Jerry Fodor calls the "dis-unity of science hypothesis" [Fodor,1974]. Yet even here my rejection of the unity of science goes much deeper than Fodor's, excluding not only inter-theoretic type identities but also rejecting token identities as well. The same thing can also be said about Karl Popper's dualism, which he describes in terms of a distinction between 1-World (physical), 2-World (mind), and 3-World (social or mind produced) systems [see Popper and Eccles,1977]. For his

dualism, like Fodor's, is a distinction only between types of events and the theories describing them, not the particular tokens of those types.

Quite generally, my view bears closest resemblance to the picture of a dualism of particular events which Jaegwon Kim has sketched in several papers [Kim,1979,1984a]. The chief difference between us is that my commitment to dualism appears much stronger than Kim's inasmuch as I reject any possible reduction of mental phenomena *via* disjunctive or species-specific properties. The precise details of my position, and the reasons for accepting a dualist picture of the mind, will be made clear in the chapters which follow.

TABLE OF CONTENTS

CHAPTER I.

TYPE PHYSICALISM AND MULTIPLE REALIZATION

In the present chapter I will defend a property dualist account of mental phenomena. My target will be the doctrine of "type physicalism," according to which mental types and properties are identical to physical types and properties. Type physicalism was made popular by the logical positivists, and was defended in one form or another by the behaviorists and the central state materialists, who at least on this score differed only on the nature of the physical types called upon to reduce mentality: overt behavior on the one hand, and processes in the central nervous system on the other [Hempel,1935; Feigl,1958; and Smart,1959].

But in the decades that followed the doctrine of type physicalism fell into disrepute. The chief cause was an argument first championed by Hilary Putnam and later reformulated by Jerry Fodor, which has come to be known as the "multiple realizability argument" [Putnam,1967; and Fodor,1974]. This argument, in roughest outline, begins with the observation that a given psychological property can be realized by any number of physically diverse systems, and ends with the conclusion that psychological properties are irreducible *vis-a-vis* the physical sciences, which is just to say that type physicalism is mistaken.

In spite of its influence, however, Putnam and Fodor's treatment of multiple realization has been subjected to a number of recent criticisms, most having yet to be answered. Hence it has become necessary to reevaluate their argument in light of these criticisms. Thus, in the present chapter I shall provide an overview of both type physicalism and the Putnam-Fodor style argument, and then examine in detail one major criticism which is particularly relevant to the phenomenon of multiple realization. Other criticisms will be taken up in the subsequent chapter. My aim will be to show that the phenomenon of multiple realization provides compelling evidence against type physicalism, and, *a fortiori*, compelling evidence in favor of a dualism of psychological and physical properties.

Preliminary Remarks on Type Physicalism

Herbert Feigl once said that: "Psychophysical identity may be identity of particulars (*this* twinge of pain with a specific cerebral event at a certain time) or of universals (pain of a certain *kind*, and a *type* of cerebral process)" [1958, p.463]. Thus, on the one theory we may identify all mental *particulars* with physical particulars, be they Cartesian substances, persons, or mental events; while on the other theory we may, in addition to this, identify all mental *properties* with physical properties, be they the attributes of sensation, properties of images, or the propositional attitude kinds. The first is called token physicalism, so named after the identity of particulars or "tokens" of mental types. The second is our doctrine of type physicalism. Type physicalism, then, is a reductionist view of the mind which can be formulated as follows:

The properties and kinds picked out or denoted by the predicates of psychological theory are identical to the properties and kinds denoted by the predicates of some physical science .

For ease of exposition I shall primarily speak about properties, since a given predicate can be said to pick out either the relevant kind or the property determined by that kind (for example, "... is a belief" can be said to denote the kind all of whose instances are beliefs or the property of being a belief). [1] But the interesting point to be made here is the restriction to *scientific* properties, those denoted by the predicates of a scientific theory.

To be specific, I shall assume that the psychological properties in question are taken from our best cognitive science, and that the range of physical properties are taken from the appropriate physical or natural science, that is, biology, chemistry, or physics. This is certainly the way matters were understood in the debate over reducing a special science to a more basic physical theory (the *Geisteswissenschaften* to the *Naturwissenschaften*), and this also fits nicely with the more traditional controversy about mind/brain identity, points which I shall return to at length.

Moreover, a quite different rationale for speaking about scientific properties is that we do not want to saddle type physicalism with the burden of identifying dubious mental entities; and psychology, being the science of the mind, is in this case our most reliable guide to what is dubious and what is not. Of course this still leaves room for disagreement over what the proper psychological kinds might be. The clinician, for example, is likely to have a different inventory of objects

than the experimental psychologist. Nonetheless there will be considerable agreement. Both will speak of beliefs, desires, images, their strength and retention in memory, their effect upon human action, and much else. Thus I assume that restricting mental phenomena in this way will allow for the stock-in-trade of philosophical discussions, the sensations and the propositional attitudes, in addition to the more theoretical constructions such as scripts, frames, mental models, and the underlying functional architecture which is deemed relevant by psychological theory. So likewise, then, should the range of physical properties be restricted to what is recognized by the appropriate physical sciences, seeing that physical science best circumscribes the legitimate bounds of the physical. [2]

Of course, this is only a provisional statement of the view. For in the next chapter we shall consider some proposals which would, in effect, extend our notion of a property beyond what is considered here. But let us assume that we know enough about the items which type physicalism holds to be identical, and move on to the question about what conditions would count towards their identity.

As a start, it is extremely difficult to give *sufficient* conditions for the identity of properties. Some, for example, believe that properties are identical if they are nomologically coextensive; still others believe the coextension must hold with a stronger notion of metaphysical necessity; while still others accept nothing less than logical or conceptual necessity. Fortunately, most of the recent discussion has centered on the nomological coextension of the properties, the chief reason being that even if this nomological equivalence is not sufficient for identity, it is

surely *necessary* . In other words, where M and P are properties denoted by the predicates of psychology and physical theory, it is enough to observe only that if type physicalism is true then it is necessary that there be the psychophysical law (x)(Mx <=> Px).

Consequently, if we have reason to doubt that the properties *are* correlated in this way, then we have reason to doubt that type physicalism is true. This is the important point, for this is just the place at which the multiple realizability argument attacks the doctrine of type physicalism.

Multiple Realization

In spite of much talk about multiple realization, philosophers have generally been content to leave this central concept unanalyzed. I shall try to remedy this situation by indicating the two most salient features of the concept. First, there is the idea that a property is realized by a number of diverse states, which I take to mean that different state types can provide lawfully sufficient conditions for the instantiation of the multiply realized property. And second, there is the idea that among this range of diverse states, there are no lawfully necessary and sufficient conditions for the instantiation of that property. This suggests, then, the following definition, where M is a multiply realized property, R the realization base, and P any property within R:

M is multiply realized with respect to a range of properties R if and only if (a) R contains some property P which is nomologically sufficient for the instantiation of M ; and yet (b) R contains no property P which is nomologically necessary and sufficient for the instantiation of M .

Clause (a) allows for the existence of one-way conditional laws of the form (x)(Px => Mx), thus giving substance to the claim that it is P in virtue of which an object realizes M. Clause (b), however, disallows the existence of any laws of the form (x)(Mx <=> Px), thus denying that there are nomological coextensions between M and any property P within the multiple realization base. This latter point is the critical one upon which the whole debate has turned, for the crux of the multiple realizability argument consists in the denial of any nomological coextensions between at least some important psychological properties and any physical properties.

So applying this to the case at hand, we let M be a given psychological property, R be the entire range of physical properties, and P any specific property within this range. The claim, then, is that *there are important psychological properties for which the concept of multiple realization , so defined , will hold .*

This claim is of great interest since, if true, it would show that a number of traditional philosophical views are mistaken. First and foremost, of course, our doctrine of type physicalism would be in error, since it is a necessary condition on the identity of properties that they be nomologically coextensive. But, in addition, any view in the philosophy of science according to which psychological theory actually reduces to physical theory by means of biconditional laws like (x)(Mx <=> Px) would be equally mistaken, at least insofar as the laws are taken to express a genuine nomological connection between the properties. [3]

Now there are roughly three reasons which have been given to accept the aforementioned claim about multiply realized psychological

properties. The first, which I call the appeal to "anomalous realization," is the denial of lawful psychophysical coextensions based on the allegedly lawless character of any generalization which contains a psychological term [Davidson,1970,1973]. The second, which I call the appeal to "macro realization," is the denial of psychophysical coextensions based upon an analogy with macro-level types, like being a bridge or a table, which have no uniform microstructures underlying their instances [see T. Nagel,1965, pp.351-352]. And the third, the argument we will canvass here, is the appeal to "functional realization" made famous by Putnam and Fodor. [4]

The Functionalist Argument

According to functionalist theory, and presupposed by standard computational models in psychology, our minds are described in the most theoretically useful way as functional mechanisms, with their mental states being described accordingly as the things which have the appropriate functional roles within this type of system. Thus, a functional property is one defined in terms of such roles, for example, the roles specified by the states of a Turing machine, a method which was once popular; or by a computer program for our cognitive system; or, what is now fashionable, by Ramsey sentences which define psychofunctional properties by their causal role in our best empirical theory of the mind [see Block,1980; and Fodor,1981]. [5]

But the important point about functional properties, however we understand them, is that they can be instantiated by objects and states of radically different composition and behavior, even to the point that

they have nothing relevant in common *physically* . To give a familiar illustration, two computers with distinct physical constitutions or "hardware" can use the same "software" or instantiate the same program. Hence the one may have computer chips made of germanium, while the other may employ the more familiar silicon chips. Furthermore, their "logic gates" may be formed in different patterns, making for quite different circuitry. Or, to dramatize the point, the second computer may have no chips at all, operating instead by vacuum tubes made of glass and wire, or by punched cards, copper needles, and wheels and pulleys, as with the ancient Babbage machine.

Whatever the precise differences in hardware, though, the two machines can nevertheless be *functionally isomorphic* insofar as they share the functions specified in their programs. Put in other terms, they can satisfy the same Ramsey predicates which define their functional properties in spite of the differences in hardware. But -- and this is the important point -- *there seems to be no chance at all that it is the same type of physical state which realizes the functional properties in each of the two systems* . A functional property in the one computer might be realized by the activation of a silicon chip, in the other by the movement of a needle through a punched card. [6] What physical property, then, is coextensive (let alone nomologically coextensive) with the functional property ? Consult all the physical sciences, there simply are no plausible candidates.

Since this point has everything to do with functional properties as such, and nothing to do with the specifics of the example, what holds for the functional states of the computer holds for the functional states

countenanced in psychology. Hence the above example can be extended to include a functionally isomorphic human being. Or, if we dislike the comparison of humans and machines, the same story can be told using only down to earth, flesh and blood creatures.

For example, our mental properties can remain constant while the neurology differs, especially when damage has occurred to one portion of the brain, and another area has taken over the psychological function [see Gardner,1985, chap.9]. Indeed, it might be that the same psychological states are realized by *nonhuman* creatures, say, those in another solar system. But given the likely differences in environment, selective pressures, and evolutionary development, it is extremely implausible to think such creatures would have anything like our own neurology.

The net result is that a type of psychological state can be realized in a seemingly infinite number of physically dissimilar ways -- by differing neurological structures, extraterrestrial X fibers, silicon chips, pulleys and wheels -- as long as these things play the appropriate role within a cognitive system. Yet among this variety of physical structures which are sufficient to realize M *there does not seem to be any physical property P which is nomologically necessary and sufficient for M* , which is precisely the claim about multiply realized psychological properties.

Indeed, take any P allegedly coextensive with M. The nature of functional properties seems to guarantee the nomic possibility of a system constructed to realize M without P, as in the case of our two computers. [7]

So much, then, for the argument. Let us now turn to the criticism.

Physical Multiple Realization

Criticism of the multiple realizability argument can be separated into two different strategies. One requires that we adjust our notion of a *property*, either by introducing the broader infinite disjunctive properties or the narrower species-specific variety, with the effect that at least some psychophysical laws will be forthcoming. This will be considered in the next chapter. The other strategy, however, leaves our notion of a property intact, and challenges the very *logic* of the argument. This is the criticism we will examine in detail here.

According to this line of response, the multiple realizability argument has been discredited by the fact that *physical properties are also multiply realized* [Kim,1972; Davidson,1973; and Wilson,1985]. Jaegwon Kim was the first to suggest that this might create a problem, and he expressed the point in this way:

> We of course should not expect to find a physical correlate for every type of mental event we commonly distinguish in daily discourse ... But the situation is hardly peculiar to mental events; we do not expect to find a microphysical basis uniquely correlating with, say, tables either. But this is not to say that tables are not physico-chemical structures or that some aspects of tables are not explicable in terms of their microphysical properties [Kim,1972, p.191; see also 1978, p.151].

Hence the common table provides us with a genuine physical type which is multiply realized by distinct microphysical properties. Of course, tables are human artifacts. Therefore it is not terribly surprising that the property of being a table lacks a microphysical coextension. But Kim also mentions the case of temperature, noting that:

> To argue that the human brain and the canine brain cannot be in the same brain state because of their different physico-chemical structure is like arguing that there can be no microphysical state underlying temperature because all kinds of objects with extremely diverse microphysical compositions can have the same temperature ... difference in material composition with respect to the kind of atoms involved, for example, does not imply difference in the mean kinetic energy of the molecules [1972, p.190; cp. also E. Nagel, 1961, p.314].

The point is that temperature is a genuine physical property, one which can even have an underlying microphysical correlate (mean kinetic energy), in spite of the fact that it is realized by various microphysical compositions, in this case, by different types of atoms.

Donald Davidson has also given a similar argument about multiply realized physical properties, only in this case applied to human action:

> It is often said, especially in recent philosophical literature, that there cannot be a physical predicate with the extension of a verb of action ... because there are so many different ways in which an action may be performed. Thus a man may greet a woman by bowing, by saying any number of things, by winking, by whistling; and each of these things may in turn be done in endless ways. The point is fatuous. The particulars that fall under a predicate always differ in endless ways, as long as there are at least two particulars. If the argument were a good one, we could show that acquiring a positive charge is not a physical event, since there are endless ways in which this may happen [1973, pp.251-252].

The idea again is that a physical event type or property like acquiring a positive charge remains undeniably physical even though, like a type of action, it can be realized in endless ways. Indeed, Davidson takes the irreducibility argument in question to be "fatuous" precisely because multiple realization is so easily attained by mental and physical alike.

Finally, Mark Wilson has recently drawn our attention to a number of multiply realized properties in the physical sciences, focusing especially on the case of temperature [Wilson,1985, pp.228-235]. What is interesting here, over and above the facts pointed out by Kim, is the observation that it is not simply that temperature is realized by various microstates, with the correlation between temperature and mean kinetic energy being preserved. Rather, that correlation generally fails, and it fails in a way that provides a striking parallel with psychological properties:

> If one reexamines the derivation in Ernest Nagel's *The Structure of Science* which seems to have propelled this particular example into the philosophical limelight, one sees that the mathematics is germane only to a classical gas and won't go through for an arbitrary system. Worse yet, there seems to be no "structural" formula of the expected type which will tell us what *temperature* should be in all substances [Wilson,1985, p.228].

We are then told that the correlation fails in quantum mechanics for dense gases at low temperatures, and that it fails in Einstein's model for solids, which is just to say that *temperature is realized by one thing in ideal gases of a certain range , by another thing in dense gases at low temperatures , by yet another thing in Einsteinian solids* , and so on [see also Sklar,1974; and Enc,1983, p.289]. But this is very much like saying that *a psychological property is realized by one thing in normal humans , by another thing in brain damaged patients , by another thing in extraterrestrials* , and the like ! Yet, as Wilson emphasizes, in spite of this analogy we are convinced that temperature is a physical property. [8]

Hence, taking all these points together, the sum of the matter seems to be this -- since we do not conclude that multiply realized physical properties are nonphysical, then, by parity of reasoning, we cannot conclude that multiply realized psychological properties are nonphysical.

The Real Lesson of Multiple Realization

But what, exactly, follows from the point that physical properties are also multiply realized ? Contrary to the above line of response, I think we can happily concede the existence of multiply realized physical properties and still maintain that the original claim has been established, namely, that psychological properties are not lawfully coextensive and hence not identical to physical properties in virtue of being multiply realized by them.

First, remember that a property M is multiply realized with respect to a range of properties R, its realization base, if and only if (a) R contains a property P which is sufficient for M, and yet (b) R contains no property P which is necessary and sufficient for M.

Now observe that it is fairly easy for a property to be multiply realized, even a physical one. Consider Davidson's case of acquiring a positive charge. The property of acquiring a positive charge is indeed multiply realized with respect to a wide range of physical properties -- simply let R be a set of physical properties which are sufficient, but not necessary, to bring about a positive charge (say, the set which includes only certain macro properties of an ignition system). Nothing of interest follows save that acquiring a positive charge is not identical to any physical property within *this particular realization base* . On the other

hand, the point is not entirely "fatuous," as Davidson claimed. For let R range over *all* the physical properties, including the microphysical property of having a distribution of protons over electrons, and acquiring a positive charge is no longer multiply realized. That is, acquiring a positive charge is identical to having a distribution of protons over electrons, and so condition (b) in our definition is not satisfied.

Or consider the case of water. Even though the property of being water is identical with that of H_2O, and admittedly physical, it is nevertheless multiply realized with respect to many physical properties -- again, restrict R to a set of properties which are sufficient, but not necessary, to insure the presence of water (say, certain atmospheric conditions, H_2O excluded, which bring about humidity). It follows that the property of being water is not identical to any property within *this particular realization base* . But let R range over *all* the physical properties, including that of H_2O, and water is no longer multiply realized.

But now consider the case of a psychological property. What is interesting here is that, unlike the previous cases, a given psychological property can be multiply realized with respect to the *entire range of physical properties !* The realization base, in other words, may contain *any* of the properties of physical science, and this, precisely because no physical property provides both a necessary and sufficient condition for its instantiation, as the functionalist argument was intended to show.

So this is why we can accept the point that physical properties, like psychological properties, are multiply realized, while at the same time maintaining that psychological properties are not identical to physical properties. The reason is that they are multiply realized with respect to

different sets of properties. The moral is that, from a philosophical point of view, *what makes matters interesting is the range of the realization base* .

How, then, does temperature stand with respect to the entire range of physical properties ? As we have seen, temperature does not have a microphysical or molecular correlate, which makes it unlike the case of water and positive charges. Nevertheless, the identification with the physical is straightfoward. For temperature is not multiply realized with respect to *all* the physical properties because temperature is itself a property of physical science (thermodynamics), and no property is multiply realized with respect to itself.

Now this answer may strike one as worse than pedantic. After all, a psychological property is not multiply realized with respect to itself either, and so in both cases we can say that the property in question is multiply realized with respect to an indefinitely wide range of physical properties *excluding itself* . So how, on the basis of this similarity, can we draw a different conclusion about them *vis-a-vis* their status as a physical property ? Or, given this similarity, why not count temperature as nonphysical too ?

But the question betrays a confusion which I take underlie this entire line of response. The confusion lies in supposing that the phenomenon of multiple realization should provide a *criterion* for what is to count as nonphysical. This, after all, was the whole point of drawing our attention to multiply realized *physical* properties. But to insist that multiple realization provide such a criterion, or that a defender of the argument must believe it to do so, is simply a mistake.

So my answer to the question: "how can we draw a different conclusion about them *vis-a-vis* their status as a physical property ?" will be that *other criteria* are used to count something as a physical property, meaning here that it is a property of physical science. And more generally, then, since other criteria are used to place an item like temperature in the domain of physical science and assign others to psychology, this leaves multiple realization to enter only afterward as a proof that the pre-categorized types are not identical.

Thus, the debate in the philosophy of psychology, taken up early on by the positivists, and more recently by Putnam and Fodor, concerned the identification of psychological properties (or predicates) with those of biology, chemistry, or physics. Clearly it was *not* multiple realization which was used to classify these properties into their respective categories. In the case of psychology, the properties and states are typically thought to be explanatory of rational activity and involved in the etiology of purposeful behavior -- criteria which are ill suited for the classification of a physical state, given that physical science seeks to explain phenomena which are not rational and which do not "behave" in the psychological sense at all.

It goes without saying, also, that since temperature does not explain rational activity, then, in spite of its being realized by various other physical states, it falls outside the category of a psychological property. Indeed, it is the different criteria I am vaguely referring to which determined the distinct sets of generalizations that we find within the sciences, crudely put, the generalizations about belief and action as opposed to mass and motion [see Fodor,1981; and Pylyshyn,1984]. In any

case, the point is that with the pre-categorized domains already in existence, multiple realization then provided a way of answering the question which would naturally arise about inter-theoretic identity.

The same thing is true in the context of the mind/body debate, as it was carried out, for example, by dualists and central state materialists. There was a certain class of things like pains and thoughts which were grouped together under the heading of "the mental." What guided the choice of such items was certain special features which they alone enjoyed, like phenomenal quality (for the sensations), intentionality (for the propositional attitudes), and on some accounts a private or direct access had by the subject of those mental states (the introspective states). The question then arose as to whether these things could be reduced to others, to be more precise, whether the mental types were identical to neurological, chemical, or basic physical types. What counted as "the physical," then, was just the things included in the physical sciences. So, roughly, a property is mental if and only if it falls under intentional, phenomenal, or introspective state types, and a property is physical if and only if it falls under biological, chemical, or basic physical types. [9]

What these historical reminders serve to show are three important things. First, it is not true that a property is mental or nonphysical if and only if it is multiply realized (and hence it is a mistake to think that an item like temperature is mental or nonphysical just because it is multiply realized). For that we use other criteria. Second, and consequently, multiple realizability should not be saddled with the work of these other criteria. And third, the only work our multiple

realizability argument can be expected to do is precisely to show that mental properties are not identical to physical properties, all understood by their proper criteria.

Put differently, multiple realization by itself provides a test for *identity* , not *nonphysicality* ; but being a test for identity, it is nevertheless adequate to show that psychological properties are not identical to physical properties. How is this accomplished ? Again, by taking the pre-categorized properties in question, the psychological and the physical, and then applying the test of multiple realization. The result is that when it is applied to the entire range of physical properties, it is exceedingly probable that the test will come out negative, as far as identity and lawful coextensiveness are concerned.

Intuitive Physical Properties

There is, I think, one final matter which needs to be resolved here. Granted, multiple realization may show that psychological properties are not identical to physical properties, where the latter are understood as the properties of physical science. All the same, perhaps this is *too narrow* a conception of the physical -- witness other intuitive physical properties like that of being a table. [10] In other words, the property of being a table is not found within physical science either, though it is presumably physical in some other sense of the term. Hence, my arguments thus far leave open the possibility that psychological properties will count as physical in some intuitively broader sense as well.

Now I concede that this is true. Psychological properties might count as physical in some intuitively broader sense. But there are three points which I think will considerably diminish whatever force this objection may have. First, as my brief historical remarks were intended to show, the conception of a physical property at play in the debate was that of being a property of physical science. The dispute in the philosophy of psychology, again, concerned the reduction of a human or special science to a physical or natural science (that is, the *Geisteswissenschaften* to the *Naturwissenschaften*); the traditional mind/body debate concerned the identity of phenomenal, intentional, and introspective state types with neurological, chemical, or basic physical types. Either way, it was physical science which circumscribed the legitimate bounds of the physical. Psychological properties, then, are irreducible *vis-a-vis* those properties, and nothing which appeals to a wider notion of the physical will detract from this point.

Second, we might try to deny the original datum which motivates this intuitively broader notion of the physical, whatever it may be, by denying that intuitive physical properties like being a table are *purely physical* . How could this be ? Perhaps because "being a table" cannot be specified without recourse to psychological notions, specifically, the *intention* that the object be used for certain *purposes* (a point which follows from the fact that there are no purely structural necessary and sufficient conditions for being a table). Now it might be thought that this point has a fairly limited application since there are other intuitive physical properties which are not artifactual and which do not involve an appeal to human intention, for example, being a rock, a tree, or even a planet. [11]

But notice that these *are* properties of physical science, of geology, biology, and astrophysics respectively; and hence they would count as physical on the criterion employed throughout, namely, being a property of physical science.

Of course, in response, one might question why psychological theory should not also count as one of these higher-level physical sciences. But our answer should now be clear. Psychological theory does not count as a physical science according to the traditional classification which was at work in the debate (it is a human or special science, not a physical or natural one; and, what lies behind this distinction, it deals with phenomenal, intentional, and introspective state types, as opposed to the other sciences). Moreover, its properties are not reducible to the properties of those physical sciences. Hence, to call psychology a "higher-level physical science" would be a victory in name only. Nothing of substance would be gained, and all the original concerns would remain -- how a human science relates to the others, how the intentional relates to the nonintentional, how a private phenomenal experience relates to the publicly accessible facts, and so on. Such concerns are not addressed by simply *renaming* psychology as a physical science.

Finally, my third point is that even if we do appeal to a broader and more intuitive notion of the physical, one which would include properties like being a table, and even if these other properties are not covertly psychological in the way suggested above, still, I doubt that psychological properties will count as physical on any intuitive scheme of classification. Why ? Because it was precisely this kind of classification which was at work in the traditional mind/body debate.

Specifically, it was an intuitive classification based upon the *prima facie* difference between the intentional, the phenomenal, and the introspective (as described in psychology), versus the nonintentional, the nonphenomenal, and the publicly accessible (as described in the physical sciences). Mental properties fall on one side. And the intuitive physical properties, like being a table, fall on the other. It therefore requires a philosophical theory to identify these intuitively distinct and separately classified items; or, as I have insisted upon, it requires the multiple realizability argument to show that they cannot be identified.

In conclusion, then, it seems that on our best and most interesting conception of what it is to be a physical property, the multiple realizability argument will suffice to show that psychological properties are not to be identified with physical properties. In particular, the mere fact that some physical properties are multiply realized by others is of no consequence. The mental and the physical will have distinct realization bases, and only in the case of the mental can a given property be realized with respect to the entire range of physical properties.

Notes to Chapter I.

[1] Those with nominalist predilections may cast our discussion in the formal mode, referring only to the alleged reducibility of the predicates. A similar argument can still be made, i.e., that multiple realization precludes any coextensivity and hence reduction of the predicates.

[2] Notice that I said "physical science" and not "basic physics." Thus, as I have construed it, type physicalism is a weaker doctrine than the traditional unity of science. For the latter requires not only that we reduce the ontology of psychology to that of a physical science, say, biology, but that we reduce, in a stepwise fashion, all such sciences to *basic physics* , thus expressing the broad philosophical view that all of reality is fundamental physical reality. As Hempel expressed it long ago: "all of the branches of science are in principle of one and the same nature; they are branches of the unitary science, physics" [1935, p.21]. Or, as Smart put it: "there is nothing over and above the entities which are postulated by physics," so that there cannot be "any irreducibly emergent laws or properties, say in biology or psychology" [1963, pp.159,160].

[3] Cp. the standard empiricist view of reduction [Nagel,1961; and Hempel,1966], although Hempel, for one, makes it clear that he is concerned only with "extensional" correlations. In any case, a different theory of reduction without biconditional bridge laws will not be relevant to our doctrine of type physicalism in any straightfoward sense at all. Robert Richardson, e.g., has recently defended a reductionist program by invoking "one-way" conditionals from physical theory to psychology [1979, pp.547-549]. But this will have no interesting ramifications for ontology [as Enc points out,1983, p.280]. More precisely, one-way conditionals provide no basis for *identifying* the properties of psychology with those of a physical science, since the latter do not even supply a necessary condition for the occurrence of the former.

[4] Because functional theories are apt to raise the philosopher's brow, it should be noted that the present argument does not require any commitment to the controversial "psycho-functional identity theory." For the latter identifies *commensense* or *folk*

psychological properties with functional properties, and this may well be false given that the commonsense states are themselves multiply realized by diverse functional states [cp. Block,1978, pp.291-293; Shoemaker,1981, pp.112-114; and recently Putnam,1988]. If this is the case, then our discussion must be restricted to the functional properties of psychology and not any common mental properties. All the same, our argument is not affected. If anything, it is enhanced. Common mental properties will be multiply realized by functional states, and they in turn by diverse physical states, making type physicalism appear doubly mistaken. The mental is, as it were, twice removed from the physical.

5 Thus, supposing Fi is the predicate variable which replaced a psychological term 'M,' we can use the Ramsey sentence to define the having of M as follows: x has M iff EF1, ... EFi, ... EFn [T(F1, ... Fi, ... Fn) & x has Fi]. And we can say that M is that property uniquely determined by the predicate on the right hand side of the formula, the having of Fi. Also, some restrictions are needed on the definitions in order to avoid the result that every property can be defined in this way. For one suggestion, see Shoemaker [1981, p.94].

6 Taken individually, the activation of a silicon chip and the movement of the copper needle through a punched card are "core realizations" of the functional property, in contrast to the "total realizations" which would include reference to other physical states of the respective machines [see Shoemaker,1981, p.97]. The point, of course, is that a physical state needs to be embedded in a proper *system* before it can realize a given functional property. In any case, multiple realization still holds, for the total realizations of a functional property can also differ in their physical instantiations. Our two computers with their radically different hardwares, e.g., provide a perfect illustration.

7 It is also worth mentioning that the point about irreducibility is not confined to the special or human sciences (the *Geisteswissenschaften*). The same thing, e.g., appears to hold for a large number of biological kinds with respect to chemistry and physics. Popular belief not withstanding, gene types are not coextensive with DNA, and the reduction bases for much of the terms from Mendelian genetics have not been found within biochemistry [see Hull, 1974]. The more general lesson is that the kind terms of any two scientific disciplines will typically "cross-classify" each other, as Fodor is wont to say, so that we cannot effect the correlations needed for reduction.

[8] Aside from this, what is perhaps the main thrust of Wilson's paper is the claim that physics employs "extremely generous mathematical methods for constructing new traits" which will guarantee that all psychological properties have a lawfully coextensive physical property [ibid., p.232]. The argument, however, is not especially clear, and the examples used appear to raise a number of different points. It suffices to say that if the physical property in question is indeed *lawfully* coextensive, then either the physical system required to instantiate the property or perhaps the newly constructed trait itself would need to be wildly disjunctive (succumbing to all the concerns addressed in chapter 2.). At least some such worry led one commentator to say that the system "must comprise most of the universe", and hence that the property in question would be physical in only a trivial and philosophically uninteresting sense [Nelson,1985, p.276].

[9] I omit reference to other physical sciences, e.g., geology, astronomy, and astrophysics, simply because no one has proposed that psychological properties be reduced to properties in their domain.

[10] I owe this point to Louis Loeb.

[11] Stephen Yablo brought this point to my attention.

CHAPTER II.

DISJUNCTIVE PROPERTIES AND THE SPECIES-SPECIFIC GAMBIT

In the first chapter I discussed the multiple realizability argument, which seemed to present a compelling case against type physicalism. I then examined in detail one particular line of response which appeals to the fact that physical properties are also multiply realized, and showed that, when rightly understood, this fact does not vitiate the original point about the irreducibility of psychological properties. In the present chapter I shall consider two more responses of a quite different nature, each one aiming to preserve the doctrine of type physicalism by adjusting our notion of a property so that the required psychophysical laws will be forthcoming. I hope to show that both kinds of response are unsuccessful.

Infinite Disjunctive Physical Properties

In a series of papers on the topic of supervenience, Jaegwon Kim has allowed for the construction of *infinite disjunctions* in order to capture multiply realized types [Kim,1978,1984b]. The basic strategy is as follows. It may be granted that a mental property M cannot be identified with a physical property P, where P is a *simple* physical property like

"having a C-fiber fire" (assuming, of course, the appropriate background conditions that the property be instantiated within a functioning human brain). The reason, as illustrated by the functionalist considerations adduced already, is that other physical properties may realize M, for example, in subjects having different types of neuron structures, extraterrestrial X-fibers, silicon chips, and so on. Hence, on the present suggestion, we are to consider instead *the disjunctive property consisting of all the physical realizations of M* , in this case, "having a C-fiber fire *or* an extraterrestrial X-fiber *or* an activated silicon chip," and so on for every physical property in the realization base for M. The net result is a perhaps infinite disjunctive property which is nomologically coextensive with M, just as type physicalism requires. [1]

Put in a different way, we need a biconditional law to express the desired correlation between mental and physical properties, like (x) (Mx <=> Px). So the strategy is to consider a more complex biconditional, like (x) (Mx <=> P_ix v P_jx v P_nx ...), where the right-hand side is now the disjunction of all the ncmically possible physical types which may realize M. Consequently, if this disjunction is thought to form a single physical property, in addition to the various physical properties denoted by each of the seperate disjuncts, then we have a method which guarantees the existence of a lawfully coextensive physical property for every mental property.

So, in short, those who defend the multiple realizability argument, on this view, simply fail to consider a wide enough range of physical properties. M is multiply realized only with respect to the entire range of *nondisjunctive* physical properties. But include within this set an

infinite disjunctive physical property, and M is no longer multiply realized. It has, on the contrary, an underlying physical correlate, though admittedly more complex than anyone might have expected.

Notice, also, that this appeal to infinite disjunctive properties does not require us to accept infinitely disjunctive *predicates* ; and neither does it require us to believe that a scientific theory must be able to use such predicates in the formulation of its laws (they would be far too complex and unwieldy for scientific purposes). Thus, we cannot rule out disjunctive *properties* merely on account of such problematic predicates [contra Fodor,1974, p.134]. Moreover, as Kim once said, we can always introduce a simple predicate by definition to be equivalent to an infinite disjunction. Hence the issue is metaphysical in nature, concerning properties, not predicates [Kim,1984b, p.172]. So, linguistic matters aside, the real question is whether such disjunctive properties are legitimate from an ontological point of view. My own opinion is that they are not, and, worse still, even if they were legitimate, they would nonetheless fail to reduce psychological properties.

Similarity and Causal Powers

The following are two standards commonly thought to decide what properties are, and each, it has been claimed, will exclude disjunctive properties. While I doubt that they provide decisive evidence against disjunctive properties, such standards do, I think, give us reason to pause before embracing the kind of disjunctive properties which Kim has proposed. Now the first standard is based on an intuitive notion of *similarity* , the one we in fact recognize, and it is this:

P is a property if and only if the sharing of P counts towards the intuitive similarity of the objects exemplifying it.

The idea is simple enough. Consider the property of "being red," and the disjunctive property "being red or round" [Teller,1983, p.58]. Red things are similar with respect to color, which is why we say they both have the property of being red in common. Similarity judgements, in other words, *ground* our talk of properties. But the things which are red *or* round are not at all like this. They seem to have nothing in common, not with respect to color (the round thing may be green) or shape (the red thing may be square). Thus, according to this line of argument, disjunctive properties are suspect given that their instances fail to meet any test of intuitive similarity.

Now what seems right about this argument is the fact that our talk of properties begins, at least initially, in the similarity we find between objects. What seems clearly wrong, however, is that our talk of properties should *end* with that similarity. Goodman predicates, for example, provide one challenge to that idea. The square red thing and the round green thing are indeed similar, one might claim, not with respect to color, or with respect to shape, but with respect to "coshape": the property of things which are either colored *or* shaped. And, indeed, Goodman predicates to one side, there are many scientific properties for which all talk of intuitive similarity is out of place.

Hence the correct thing to say, I think, is that the similarity which determines objective propertyhood is a "modified similarity," as Quine put it, one which arises out of scientific concerns [see Quine,1969, esp. pp.128-129]. In other words, it is our intuitive notion of similarity

tempered by the ongoing practice of science.

This appeal to a modified similarity has a number of consequences. First, it resolves the problem created by Goodman predicates, since not only are Goodman predicates unintuitive, but they fail to be entrenched in our scientific theories. Second, and this is really the important point for our present concerns, a scientifically modified similarity will resolve the controversy which surrounds the infinite or indefinite disjunctive physical properties that Kim has proposed. For, like their Goodmanian counterparts, terms which pick out these disjunctive properties also fail to be entrenched in our scientific theories. (Whether we should make a special allowance in this case, given the utility of such disjunctive properties for inter-theoretic reduction, is a question I shall return to later on when I discuss their explanatory value.)

In any case, before we leave the topic of similarity, let me add an important word of caution. This appeal to similarity need not be taken as a blanket rejection of all disjunctive properties. Consider, for the sake of argument, the identification of a genus with a disjunction of all the species falling under it, "being an animal" with "being a human or rabbit or duck," and so on. In this case the things which satisfy the disjuncts do indeed have something relevant in common, something which makes them animals, and it is a similarity which could, I assume, be specified in biological terms. This is why we might be tempted to say that the disjunction picks out a physical property in spite of its disjunctiveness. But the disjunction proposed by the type physicalist is not at all like this. The things which underlie a given mental state -- the various neuron structures, X fibers, silicon chips, pulleys and wheels -- have

nothing in common physically, no similarity with respect to physical science which makes them realizations of mentality, and that is the crucial difference.

Can the type physicalist appeal to the *mental property*, realized by the disparate elements, as the relevant feature which the physical disjuncts have in common ? No, for then we would not be giving a reduction of the mental property solely in terms of the properties of a physical science. Moreover, this is not what we would say in the genus/species case. We would not say "all the humans, ducks, rabbits, and the like, have the property of 'being an animal' in common. That is what makes them animals." Rather, we would try to specify certain biological facts which the creatures share, perhaps similarities in DNA structure, or, as a more traditional classification has it, a similarity defined in terms of the propensity for spontaneous movement along with a capacity to respond to external stimuli. So, in short, the trouble with the various realizations of mentality is that they simply have nothing relevant in common physically, save that they are realizations of mind.

Now the second standard commonly thought to determine the nature of properties is based on the notion that properties confer *causal powers* on the objects which have them, and it is this:

P is a property if and only if the having of P makes a contribution to the causal power of the objects exemplifying it.

The idea, in other words, is that properties have a certain work to do in our overall scheme of things, and at least some of that work is not only to ground the "objective resemblances" of objects in our ontology, but also to explain their "causal powers" [e.g., Armstrong,1978a; and

Shoemaker,1980]. But disjunctive properties seem to do no such thing. That is, they do not seem to play any role in the causal relations between things, and that, arguably, is what distinguishes real properties in the world from mere concepts about them. For example, in the case where an object's having F brings about an effect, it is the property F which is causally efficacious, not the disjunctive property F v G [Armstrong, 1978b, p.20; see also Fodor,1986, p.11].

But, again, a word of caution is in order. This appeal to causal powers need not be taken as a blanket rejection of disjunctive properties either, though it probably has been so taken. For example, consider a case of causal overdetermination, say, turning on two light switches at the same time, each being sufficient to light the bulb. Suppose the one switch being turned on is an event of type F, the other of type G, and the effect in question E. Then someone might argue that the best candidate for being the cause of the E type event is the disjunctive event F v G. [2] That is to say, it is not the event of type F, since the counterfactual "had the F not occurred, the E would not have occurred" does not hold, and this, due to the sufficiency of G. And it is not the event of type G either, and for the very same reason. Yet it is not the event of type F & G, for that is much too strong, each conjunct being individually sufficient for E. Thus, F v G appears tailor made for the job at hand.

But even if we accept this way of handling overdetermination, [3] I think there is still an important difference between this appeal to disjunctive properties and the infinite disjunctions required by type physicalism. For notice that the overdetermination case requires the *entire disjunction to be relevant in the circumstances* . What I mean is

that property F is exemplified in the circumstances and presumably contributing to the production of E, and so is G. Moreover, had it not been for the exemplification of both properties, the appeal to the disjunction F v G would have been superfluous, with one of the disjuncts being causally idle.

Hence, call the kind of case in which both properties are exemplified an instance of *inclusively exemplified properties* . We can then say that the important difference lies in the fact that the disjunctions proposed by the type physicalist are *exclusively exemplified properties* . For here only *one* of the disjuncts is typically exemplified at any one time, the rest being purely counterfactual realizations of the mental property. Consequently, these other disjuncts are not causally relevant in the circumstances simply because they are not exemplified at all.

Both points, I think, could be summed up by saying that the real properties in the world are those which have *explanatory value* . For both objective similarities and causal relations are things which can enter into our explanations of the world order. This, I suspect, is why the proposed infinite disjunctions appear *ad hoc* , [4] since they have no part in our common conceptual scheme, and no part in the explanations of any serious science. Of course, one might try to deny this. One might say that infinite disjunctive properties *are* required by science, specifically, in order to reduce disciplines like psychology to those of physical science. But this surely would be *ad hoc* . For the proposed infinite disjunctions have *no independent role in the scientific theory* , and absolutely *no purpose other than to reduce the unwanted entities* -- unwanted, I might add, not from the perspective of any scientific discipline, especially not

from psychology, but from certain purely philosophical motives.

Compare the case of water and H_2O or gold and the things having a structure with the atomic #79. The chemical properties and microstructures are not introduced with the sole purpose of reducing the respective macro substances. They have their *own* explanatory work to do, for instance, to explain the effects of chemical interactions. The proposed infinite disjunctions, on the other hand, have no such explanatory value. Ironically, they show themselves to be genuine nomological danglers. [5]

Now one might question whether this appeal to explanatory value is actually critical to the debate. [6] For if our concern is one of ontology, a concern with what properties actually exist independently of our theories about the world, then matters of explanation may seem irrelevant. Put in a different way, the arguments given thus far might refute an "explanatory reduction," but not a more cautious "ontological reduction" [cp. Rosenberg,1978, p.382].

In response, it should first be observed that the debate over the reduction of psychology arose with explanatory properties in mind, that is, with *scientific* properties. Thus the appeal to infinite disjunctions can hardly be said to address such concerns. But more importantly, I do not think we can radically divorce ontology from our explanatory concerns, as the defender of infinite disjunctive properties would have us to do, inasmuch as our best explanatory schemes are the most reliable guide to what exists. Following Quine, we could say that our posits should not go beyond what is required for the ongoing practice of science, and yet infinite disjunctive properties are simply not counted

as among those posits.

The Irrelevance of Physical Disjunctions

Finally, even if disjunctive properties are not ontologically objectionable, in the way I have outlined, the appeal to infinitely disjunctive physical kinds will nonetheless fail to reduce psychological properties. I shall mention one consideration, which is that the appeal seems to miss the point which originally motivated functionalist psychology. For the idea was that the essence of a psychological state type is to be specified by its functional role within an information processing system. But, as many pointed out, this places no constraint on the *kind* of mechanism that realizes it. Hence it places no constraint on the *nature* of the information processing system, be it physical, soulish, or even the mind of God, except that it be such as to process information [see Putnam,1967, pp.435-436; Block,1980, pp.177-181; Boyd,1980, pp.92-97; and Shoemaker,1981, p.98].

This being so, then even if a given mental type is in fact realized by only physical mechanisms, nothing precludes the possibility of a nonphysical realization. But this means either that the proposed infinite disjunction of physical kinds is not large enough because it will not support counterfactuals for those worlds where there is a nonphysical realization, or else it means that the infinite disjunction has a nonphysical realization as one of its disjuncts, in which case physicalism has been abandoned. Either way, type physicalism must be rejected.

I suspect, however, that an austere physicalist may want to say that while these nonphysical realizations are conceptually or logically possible, they are nonetheless *nomologically* or *metaphysically impossible* [cp. Horgan's definition of "physically accessible" worlds, 1982, p.35]. But this appears to be merely an act of legislation on the part of the physicalist, an assumption which we are in no way forced to accept. Indeed, and possible worlds aside, there is nothing wrong with having nonphysical realizations of nonphysical properties, even in the actual world, as the case mathematical objects will attest. And even if no irreducibles of a more robust variety actually exist in our world -- God, Cartesian souls, entelechies, and the like -- our world might be such that, in David Lewis' words, "spirits are absent but not outlawed" [1983, p.363]. [7]

In the very least, given that the nature of functional states places no constraint on the composition of the system containing them, the burden of proof is on the physicalist to show why it is impossible, nomically or otherwise, for them to have nonphysical realizations. My conclusion, then, is that either infinite disjunctive properties are ontologically objectionable, or else, if they do exist, their employment by the physicalist will not succeed in reducing psychological properties.

Species-Specific Laws

The last criticism of the multiple realizability argument I want to discuss was also suggested by Kim, and it has become quite popular in some circles. The interesting idea here is that, rather than appeal to a wider notion of a property via the controversial infinite disjunctions, we

are to consider a more narrow class of "species-specific" properties. As Kim put it:

> Let us assume that the brain correlate of pain is species-dependent, so that we have generalizations like "Humans are in pain just in case they are in brain state A," "Canines are in pain just in case they are in brain state B," and so on. Theses species-dependent correlations do not of course warrent the species-independent blanket identification of pains with a "single" brain state ... But they clearly do warrent -- at least they are not inconsistent with -- the identification of *human pains* with *human brain state A* , *canine pains* with *canine brain state B* , and so on. That is to say, *species-specific correlations warrant species-specific identities* [Kim,1972, p.190].

Thus, while the multiple realizability argument may show the impossibility of *completely general* laws connecting the mental and physical, this does not preclude *species-specific* laws which reduce species-specific properties, for example, "human pain" versus "canine pain," and the like.

David Lewis has taken a similar position by defining a mental type like "pain" as being a state which occupies a certain causal role, a role that he relativizes to a "population," which he says is "a natural kind -- a species, perhaps" [Lewis,1978, pp.219-220]. And Berent Enc has also followed Kim in this regard, finding support in the reduction of temperature to the kinetic theory of gases. As we discussed in the last chapter, temperature is multiply realized with respect to various microstructural properties, being realized by one thing in the case of gases and something else again in the case of solids [Enc,1983, p.289; cp. Wilson,1985, p.228]. Enc goes on to say:

> However, the interesting aspect of the thermodynamic example is that all temperatures over which the gas laws have jurisdiction are identical with mean kinetic energies. To put it more bluntly, the sceince in which the gas laws are formulated, like psychology, is a special science, and special sciences are restricted to a domain of objects; the objects of the gas laws are gases. A reduction may succeed in identifying a property P of gases with some property Q of its constituents, and yet the property P of some objects which fail *outside* the domain of that special science may not be realized by Q ... If this is the correct account of the reductive relation between a special science and a more basic one, then when we examine the possibility of reducing psychology to neurology, we must allow for the fact that the proper objects of psychology are human beings [Enc, ibid., pp.289-290].

Of course, Enc is speaking somewhat loosely here. For in the thermodynamic example, the special science cannot consist of the gas laws *simpliciter*, since temperature is not identical with mean kinetic energy in the case of dense gases at low temperatures [see Wilson, ibid.] Hence the laws will not range over *all* the gases, and the domain of objects is consequently restricted to a subset of the gases -- perhaps making it a peculiar special science. Moreover, it is not "temperature," as Enc says, which is identical to mean kinetic energy even in this restricted set of gases, just as it is not "pain" which is identical to a type of human neurology. Rather, it is "temperature *for some gases* " and "pain *for humans* " which is being identified. In any case, the species-specific analogue is clear: relativize common mental states to a restricted domain of objects, as we relativize temperature to a restricted set of gases, and the reduction of properties will follow.

Finally, D.M. Armstrong has also adopted this species-specific gambit. In a recent debate with Norman Malcom, Armstrong observes that what

plays the causal role of pain may differ for different sorts of creatures, which is just the point about multiple realization. He then says:

> It may be granted, for the reason just discussed, that it is impossible to identify the type *pain* with a certain neurophysiologicalcal process. But what about the more narrowly conceived type: *pain in human beings* ? It is quite plausible that it can be identified with some single sort of neurophysiological process [Armstrong, 1984, p.162].

To summarize, then, the idea is to no longer search for the dubious psychophysical law (x) (Mx <=> Px), one which employs the more general psychological properties, but to find instead the law (x) (M_Sx <=> P_Sx), where 'M_S' and 'P_S' are the more narrow species-specific properties to be identified.

So the important question is whether psychological theory can allow for species-specific properties, and employ them to the exclusion of the more general and irreducible properties which we discussed in connection with the multiple realizability argument. *Prima facie* it would seem not. For the multiple realizability argument was based upon the descriptive practice of psychology, being grounded in the way psychological theory classifies its objects into functional kinds. Moreover, it would seem to be a desideratum of *any* scientific psychology that it be able to capture interesting generalizations which obtain across species. Yet such generalizations, if there be any, must in principle lie outside the purview of a species-specific psychology.

Of course, the functionalist must be careful here as well. Specifically, the more fine-grained the functional description may be,

the less likely it is that another creature will, in fact, satisfy the description. This is one reason to adopt Fodor's more modest version of functionalism, for he only wants functional definitions for the propositional attitudes and not their content, that is, functional definitions for *belief* as opposed to *belief that P* [see Fodor,1987, pp.69-70].

Thus, where "being a belief" is defined in terms of the basic functional architecture necesssary to support beliefs, roughly, by whatever it takes so that the system can be said to have a belief box, and where "being a belief that P" is defined in terms of its conceptual role with respect to all other semantically relevant beliefs Q, R, and S, then clearly the first definition can apply to a much wider range of cognitive systems than the second, namely, those believing creatures which have the same basic functional architecture but who happen to lack any of the particular beliefs among P, Q, R, and S.

But however comprehensive we might want our functional definitions to be, there are at least three considerations which might lead one to think that the restricted species-specific properties are indeed compatible with a scientific psychology, and required by it, even as it is practiced today.

Evidence for Species-Specific Properties

First, there is the fact that there are cases of species-specific psychological laws. Rats can be conditioned to avoid food by taste, but not by sight. Yet birds are exactly the opposite. They can be conditioned to avoid food by sight, but not by taste. [8] It might be tempting, then, to

infer the existence of species-specific properties from these *psychological* laws, thus making it plausible to think that species-specific *psychophysical* laws will be forthcoming.

Second, one might think that the phenomenon of multiple realization actually *presupposes* species-specific correlations. Kim has given an argument to this effect:

> In fact, it is a tacit assumption of the [multiple realizability] argument that there are *species-specific* psychophysical laws ... Indeed, the very notion of 'physical realization' of pain seems to presuppose the existence of nomological connections, *within each species* , between pain and some underlying neural process. If there were no such nomological link, in what sense does *this* neural state, and not some other one, 'realize' pain ? And how would we know that it, and not some other state, is the physical realization of pain for this species ? [Kim,1982, p.55; cp.1979, pp.33-34].

Third, one might support species-specific correlations by appeal to a model of theory reduction in the philosophy of science. The idea is that scientific theories are often described as being "corrected under reduction." The Newtonian reduction of Kepler's first law of planetary motion is a case in point. What follows deductively from Newton's law of gravity is not that planets travel in perfect ellipses around the sun, as Kepler thought, but that they deviate from those paths, if ever so slightly, due to the gravitational attraction of the other planets. Nevertheless, there is such a close resemblance between Kepler's first law and the corrected law with approximate ellipses that we say Newton's law reduces the former by reducing the latter.

Similarly, it might be suggested that functional psychology stands to be corrected under reduction by restricting a domain of psychological entities to human beings, thus generating species-specific laws which are more plausible candidates for deduction from a physical theory. In this way psychology might be reduced to physical theory in virtue of the fact that *human* psychology, and all the other species-specific psychologies, are so reduced. In Kim's words, reduction will proceed in terms of these more "local physical coextensions" [1984, p.173].

Failure of the Species-Specific Gambit

Let me respond, then, to these considerations in support of species-specific psychophysical laws, beginning with Kim's view that they are actually presupposed by considerations about multiple realization. That view, as I understand it, requires that a species-specific law ground the claim that one state realizes another. Now it is true that a species-specific law would be *sufficient* to justify the claim that some physical state is the realization of a psychological property. But I do not think it is *necessary* . A one-way conditional law which employs the more general properties like (x)(Px => Mx) would seem to do the job equally well, as long as P is one of the many physical properties which may bring about the multiply realized M. So the species- specific law (x) (M_Sx <=> P_Sx) is not presupposed.

Turning now to the point about scientific theories being corrected under reduction, and the claim that functional theory stands to be corrected by a species-specific psychology, we would do well to take a closer look at the kind of theory reduction in question.

As a start, Kenneth Schaffner has formulated a model which is intended to capture the reduction illustrated by the paradigm of Kepler and Newton's laws. On this view, a theory of some primary science T_1 (like Newton's) reduces a theory of a secondary science T_2 (like Kepler's) when there is yet another theory T_2^* (the one that approximates Kepler's) such that: (a) T_2^* is *deducible* from T_1 by means of biconditional bridge laws; (b) T_2^* *corrects* T_2 in the sense that it leads to more accurate predictions than T_2; (c) T_1 *explains* T_2 in the non-formal sense that T_1 has T_2^* as a deductive consequence, a theory which bears "close similarity" to the original T_2 and whose predictions are "very close" to those of T_2; and finally, (d) the relation between T_2 and T_2^* is accordingly one of "*strong analogy*" [Schaffner,1967, p.144].

To apply this to the case of psychology, then, and focusing on the first two conditions, the claim is that our psychological theory T_2, with its functional descriptions that can range over physically quite different systems, is to be corrected under reduction by a species-specific psychology T_2^*, which is, I assume, a conjunction of human psychology and martian psychology, and the like, such that: (a) species-specific psychology is *deducible* from a physical science by means of biconditional bridge laws; (b) species-specific psychology *corrects* functional psychology by leading to more accurate predictions than the original theory; and so on for conditions (c) and (d) which require that a strong analogy obtain between the two theories.

Now let me say at the outset that the following points are not essentially tied to the particular model of reduction given here. [9] Indeed, they would seem to apply to any model of species-specific

reduction relevant to our topic. Thus, it should become evident, I hope, that functional psychology does not stand to be corrected under reduction in any way like this.

Consider first condition (b), that the species-specific psychology correct functional psychology by having more accurate predictions. This, I think, is simply false. Of course this claim, like its denial, is tenuous since functional psychology is in its infancy and species-specific psychology is nonexistent. Nevertheless, we can say this much. If the species-specific psychology employs *functional* descriptions, specifically, if the descriptions formulated with particular species in mind are cast in functional terms, then the two psychologies in question will have equivalent predictions (as long as we do not assume that functional psychology employ the fine-grained descriptions mentioned earlier, or that it be committed to the idea that all cognitive systems instantiate the same program).

If, however, the species-specific psychology does not employ functional descriptions, but, say, a vocabulary reducible to the respective neurologies, then while it may lead to accurate predictions of some sort, it will not, I think, have *more* accurate predictions simply because the theories will be about quite different things. What I have in mind is the commonplace observation that, for instance, neurological theory would count some behavior as distinct which psychological theory considers the same. Signing a check and paying cash are physically distinct, to use Fodor's example, yet psychological theory might count them the same if the subject would in each case be exhibiting the same behavior, that is, paying his bills [see esp. Pylyshyn,1984, chaps.1 and 2].

But then the neurophysically reducible theory may have accurate predictions about the movement of hands with pens in them, as opposed to the movements towards wallets, or whatever the physical behavior may be, *but the psychological theory can at the same time have accurate predictions about people paying their bills* . The two theories simply do not compete for the same logical space.

Let us turn, then, to what is perhaps the most important condition, namely, (a) that species-specific psychology be deducible from a physical theory. I think this is also false. That is to say, the appeal to species- specific properties and laws is ultimately unsuccessful because *the phenomenon of multiple realizability can occur within a species* . This is sometimes put by saying that there is considerable *plasticity* in how the brain realizes psychological functions. To take an extreme case, this would be true if Karl Lashley's doctrine of neurological "equipotentiality" were correct, that is, if psychological functions can be subserved by any number of brain structures [see Gardner,1985, chap. 9; cp. Block and Fodor,1972, p.238]. The point is that they are all states of a particular species, and a very large number indeed.

Now Lashley's doctrine, as it turns out, is probably incorrect in the very broad form in which he gave it. Many psychological functions appear securely localized. For example, the most frequently cited case is linguistic representation in the left hemisphere [Penfield and Roberts, 1959]. But even this is not universal across the human species, because a minority has speech located in both centers [Zangwill,1960; Sperry, 1974], and because children have some bilateral representation until

about the age of five years [Popper and Eccles,1977, p.313]. Indeed, the point about plasticity is best illustrated in cases where a transfer of cognitive function has occurred because of some neurophysical damage. As Gardner summarizes our present state of knowledge:

> [I]mpressive evidence continues to accumulate documenting the resilience and plasticity in the nervous system, particularly during the early stages of development. At such times, even organisms deprived of the usual neuro-anatomical structures are able to adapt and to carry out requisite functions, sometimes without incurring excessive costs [1985, p.278].

Hence, the alleged species-specific correlations will need to countenance large disjunctions of neurophysical types, making then subject to the previous concerns over disjunctive properties, and rendering the overall strategy unsuccessful as an attempt to find simpler and more acceptable psychophysical laws. Of course, those who invoke species-specific properties are not unaware of this problem. After appealing to the property of *pain in human beings* , Armstrong says:

> And if even that identification turns out to be too optimistic, it will presumably be possible to find still narrowly conceived sub-types: pain in human beings of the sort X, in human beings of the sort Y, ... and so on, where the identification can finally be effected. For, after all, the idea that the physiological nature of pain in human beings changes from occasion to occasion, or even from person to person, seems truly bizarre, although it may be a logical possibility [Armstrong,1984, p.162].

In response, one is tempted to say, first, that what Armstrong finds as "truly bizarre" and a mere "logical possiblity" is precisely the facts to be derived from the empirical data concerning individual differences in

how psychological functions are realized (if "pain" is securely localized, again, there are a host of other psychological states for which the point about plasticity holds). But, more than that, just what *are* the "more narrowly conceived sub-types" to which Armstrong refers ? This is no place for vague gestures at what might solve the problem. The plasticity of the brain presents a formidable challenge to anyone who adopts this kind of strategy. Yet we are not given the slightest hint at what the more narrow sub-types might be.

We cannot, for example, appeal to the property "pain in human beings *who have neuron structure X* ," the neurons which happen to realize pain in that particular group, for this would be a surreptitious introduction of the same physical property which is supposed to reduce the psychological property.

A different kind of response, however, and one which I think is quite natural given the overall strategy, might be this. In light of the individual differences within a species, we relativize the mental property M to an individual S, creating the property M_S to be identified with some physical property. Unfortunately this will not do. For, as we have seen, the realization of a mental property may change within an individual over time due to brain injury and the like. Thus, in light of these differences in S over time, we relativize M_S to a particular time t, creating the property M_{St}, again to be identified with a physical property. But now the view has become, for all intents and purposes, indistinguishable from a *token identity thesis* . In other words, relativizing M to S at a time t seems very much like talking about *S having M at t* , which philosophers have recognized as a dated particular,

that is, a token structural event [see Kim,1969; and Lombard,1986].

In the very least, until those who defend species-specific properties give us an indication of what the plausible sub-type might be, then we cannot say that the species-specific strategy avoids the problem created by the plasticity of the brain. Moreover, this plasticity of the brain *vis-a-vis* the psychological functions also provides an answer to the first consideration mentioned earlier, namely, the temptation to infer the existence of species-specific psychophysical laws from the species-specific psychological laws (like those for rats and birds). For we can say that if there were species- specific psychological laws for humans, the properties involved would be multiply realized by diverse human neurophysical states, preventing the desired psychophysical correlations.

Finally, and this is a point worth emphasizing, even if there were no such plasiticity, no individual differences for the human species, nevertheless, given that these psychological states restricted to humans are functionally specified, then *the very same multiple realizability argument could be run in terms of these functional properties* . For assume that there is a Ramsey predicate specifying a human psychological function, and that it is satisfied by all and only human beings. However, given the nature of functional states, it is merely a *de facto* truth that only human beings satisfy the description.

This is, I think, a very real possibility. To illustrate, consider the recent "connectionist" models of cognition. As Patricia Churchland has described them, such models typically have three architectural elements: processing units, connections between those units, and weights, which

are the differential strengths of connection between them [Churchland, 1987, p.550]. The flow of information is then viewed as a pattern of activation over such a network. Now these models represent a "bottom-up" strategy, as Dennett would say, based upon the actual workings of the brain. Specifically, the processing units correspond to neurons, and the connections and weights correspond to the differing firing rates between those neurons. And because of this it might be plausible to assume that, as a matter of fact, the human brain alone satisfies the model, and consequently that whatever laws obtain between the states satisfying this model are species-specific in just this sense.

All the same, *the model could be satisfied by systems having no neurons* , perhaps if another creature evolved in the right way, or if we actually built a functionally isomorphic system to do so. For example, the processing units could be small photo-sensitive machines which emit and receive light, the connections and weights being the various frequencies of light transmitted between them. So, again, it is merely a *de facto* truth, if it is true at all, that only certain species will satisfy a psychological description. Put differently, it is one thing to say that a psychological law applies to a specific species; it is quite another to say that it is necessary that the law apply *only* to that species. The inference from species-specific psychological laws to species-specific psychophysical laws is therefore seen to be fallacious.

In conclusion, then, we have examined several of the more important objections to the multiple realizability argument, and none, in my view, seem especially plausible. Some, perhaps, rest on misunderstandings

about what the multiple realizability argument requires. Others depend upon objectionable philosophical moves, and unnatural restrictions on the scope of psychological generalizations. The multiple realizability argument, on the other hand, serves to explain in a convincing fashion why psychological properties should not be expected to reduce to physical properties. Commitment to a dualism of properties, therefore, seems to be the most reasonable approach in matters of the mind.

Notes to Chapter II.

[1] William Taschek has pointed out to me that if we are interested in mere nomological coextensivity, then the required disjunction may very well not be infinite, but very large. Hence, any argument against the use of such properties for reductive purposes must not hinge on the fact that they are *infinitely* disjunctive. Fortunately, the first two arguments in this chapter will be directed merely at the *disjunctive* aspect of these properties, regardless of whether they are infinite or just indefinitely large; and the third argument maintains that even an infinite physical disjunction will not reduce the mental property.

[2] I owe this point to Stephen Yablo.

[3] My own view, following Martin Bunzl [1979], is that such cases are typically underdescribed, and that the genuine possibilities turn out to be cases of causal *preemption* , for which the cause is at least intuitively clear. Also, I think the existence of disjunctive properties will actually *proliferate* cases of overdetermination, not resolve them. For suppose an event of type F causes an event of type E. Given disjunctive properties, and given that having F entails having F v ... , what prevents one from invoking the events of type F v G, or F v G v H, and so on indefinitely, thus making the effect overdetermined ? Of course, properties like G must be causally relevant for E type events (even though, *ex hypothesi* , they were not active in this case), otherwise, e.g., a regularity theorist could exclude F v G on grounds that the covering law (F v G) => E is false due to the falsity of G => E.

[4] It is also interesting to note that Kim had previously expressed some reservation about disjunctive properties: "Even if we disallow the *ad hoc* creation of new states by forming arbitrary disjunctions ..." [1972, p.190]. Indeed, Kim has sometimes held that events which enter into different explanations are on that account different events, a view which entails that events, along with their constitutive *properties* , are identical if and only if they are in some sense *explanatorily equivalent* [see Kim,1966, p.232; 1969, p.201; also Lombard,1986, pp.50-54]. Properties, then, must have explanatory value on this view.

5 This, of course, is a backhanded reference to Smart's claim that irreducible mental properties would be nomological danglers. However, at this point I should mention a more subtle response which the type physicalist could make. For the type physicalist could say that if mental properties *are* disjunctive physical properties, then any explanatory value had by the mental properties will *ipso facto* be had by the disjunctive physical properties. True enough. But I take it that this response would be question-begging in the present context, since the proposed identification is precisely what is at issue. Hence, as far as their explanatory value is concerned, we should remain neutral with regard to the identity question. Moreover, even if i am wrong about this, and the explanatory value of a mental property is allowed to accrue to the disjunctive physical property, still, this kind of explanatory value remains suspect. For, as I argued in the text, the disjunctive physical property has no explanatory value *within physical theory* (its explanatory power would derive solely from the role it plays within psychology).

6 As William Taschek suggested to me.

7 It should be noted, however, that if we accept the nomic possibility of nonphysical realizations, this will affect our view of psychophysical supervenience (a doctrine I discuss in chapter 5). E.g., in order to account for the possibility of divine and soulish thoughts, and at the same time to maintain a significant version of psychophysical supervenience, we could restrict the supervenience claim to spatio-temporally extended beings.

8 I owe this example to Ed Smith. Also, even if all known species-specific laws were of the baser "conditioning" variety, they do suggest the possibility of such laws for the higher cognitive functions.

9 The justification for using Schaffner's model is that other accounts of how one theory may correct another, e.g., the one proposed by Kemeny and Oppenheim [1956], will simply not be relvant to our doctrine of type physicalism in any straightfoward way. The reason is that type physicalism requires *biconditional* laws, as in the standard empiricist model of reduction [e.g., Nagel,1961; and Hempel,1966], which Schaffner wants to preserve. Yet the Kemeny Oppenheim model does not require *any* laws connecting the theories of a primary and secondary science. Rather, the former is only required to have roughly the same

observational predictions as the reduced theory. The result is that the secondary science is not *reduced to* but simply *replaced by* the primary science, much in the way that phlogiston was replaced by Lavoiser's oxidation theory. In a word, the replacement model is relevant to eliminativism, not type physicalism.

CHAPTER III.

TOKEN PHYSICALISM AND THE NATURE OF MENTAL EVENTS

The kind of physicalism which has become the dominate view among philosophers is called "token physicalism." Remembering the distinction between the identity of *universals* , *attributes* , or *properties* , and the identity of *particulars* , *substances* , or *objects* , that is, between type and token identities, the present doctrine is a claim only about the latter. On this view, all mental particulars must be identical to physical particulars even if the kinds to which those particulars belong or the properties which they have in common are not reducible to physical kinds and properties. This may include the identity of persons with material bodies, minds with brains, and, more typically, the identity of mental events with neurophysical events. For example, a particular thought of Vienna must be identical to a particular neurological event even in the absence of the general type identity between thoughts of Vienna and neurological states of that kind [see Nagel,1965; Davidson,1970; and Fodor,1974].

As an historical note, although central state materialists like Feigl and Smart had noted the difference between type and token identities, and had defended the stronger type physicalist view, the earliest hint of

a purely token identity theory came out of an exchange between Plantinga and Putnam on Turing machine functionalism. Acknowledging Putnam's argument against type identities, Plantinga said:

> I doubt, however, that the identity theorist would wish to dispute Putnam's conclusion; for I am inclined to think that when he says being in pain is really being in a certain neurological state S, the identity theorist does not mean to assert the identity of any universals at all. What he means to assert is that every instance of the universal being in pain is contingently identical with some instance of the universal possessing neurological state S... [Plantinga,1962, p.203].

I take it that what Plantinga called the "universal" is the mental type, and what he called the "instance" is the mental token, which later became the token events, states, and processes explicitly discussed in the literature. However that may be, I suspect that philosophers have been attracted to the idea of token identities mainly because it would provide a physicalistic view of the mind which is significant, but significantly different than what is provided by the more austere doctrine of type physicalism which Putnam and others had rejected. Indeed, I suspect that many physicalists retreated to a less objectionable claim about mental tokens precisely because of the failure of the type identity theory.

Unfortunately, I think this retreat was made in haste. I shall argue that token physicalism is metaphysically unsound, resting, as it does, on an implausible theory of events. Hence the following will be an argument against token physicalism based upon the metaphysics of events. In the present chapter I will establish the connection between the individuation of events and token physicalism, specifically by showing

how the two leading theories of events lead to different conclusions about the status of token identities. Then, in the next chapter, I shall argue against token physicalism by attempting to rule out that theory which allows for token identities. The result, I think, is that we should countenance not only a dualism of properties, but a dualism of particulars as well. The ontology of psychological theory turns out to be irreducible across metaphysical categories.

Theories of Events

All of the discussion concerning the individuation of events has so far centered around two or three of the leading rival theories. I will follow this practice here, knowing that this may leave out some other promising theory. Nevertheless, there is some justification for limiting the scope of inquiry in this way, namely, the leading theories are typically those which are thought to be the most plausible candidates for truth. The discussion will not be unduly limited, however, since I will also consider modifications and extensions of the leading theories which have been suggested in the literature.

I should also emphasize at the outset that we are concerned with a *metaphysical* issue about individuation, not a *semantic* issue about the correct truth conditions, or the proper logical form, of event describing sentences. For example, much has been written about Davidson's linguistic theory, according to which we parse sentences of action like "Jones buttered the toast in the bathroom at midnight" into the form "Ex [Buttered(toast, Jones, x) & In(bathroom, x) & At(midnight, x)]," meaning that there is an event x such that x is a buttering of the toast by Jones, x

is in the bathroom, and so on for each additional clause [Davidson,1967a]. So each action sentence contains an implicit quantification over events, events which are described in various ways by the adverbial clauses now turned into predicates for events.

But Davidson also has a metaphysical theory about how events are to be individuated, which is a quite different matter, since there are many theories of events on the market, any one of which could lay claim to providing the domain of discourse for Davidson's quantified sentences [see esp. Kim,1976, pp.163-164]. [1] This is not to deny that Davidson's linguistic theory, if correct, would provide *some* information about the nature of events. For, given his semantics, the events are described, dated, and localized, making them appear as concrete particulars. Thus, a theory of events as *universals* or purely *abstract states of affairs* would be effectively ruled out [e.g., as in Chisholm,1970 and 1971].

Moreover, whatever we think about abstract states of affairs, and whatever purpose they may serve (perhaps, as Chisholm suggests, to account for our talk of "recurrence"), it seems that we must talk about *particular* events if we are to give an account of the interactions and alterations which occur between concrete objects [see Thalberg,1980]. [2] Thus, we need more than a semantic theory for event sentences, and in particular, we need a criterion of identity and individuation for the class of particular events.

Now Davidson's first suggestion, which I will just mention and criticize in passing, was to individuate events in terms of their *causes* and *effects* [Davidson,1969; anticipated by Nagel,1965, pp.346-47]. As he put it, where x and y range over events:

(DC) (x)(y) (x=y iff (z) (z caused x <=> z caused y)
and (z) (x caused z <=> y caused z)).

This criterion, it has been charged, has an air of circularity about it. But Davidson thought it was not formally circular since no identity sign appears in the definiens, or the right hand side of the main biconditional. Recently, however, Davidson has recanted [Davidson,1985a, p.175]. As Quine made him aware, the right hand side does quantify over events, and this makes sense only if events are *already* individuated [Quine,1985, p.166; the same criticism was made much earlier by N.L. Wilson,1974; and discussed in Brand,1976, p.138]. In simplest terms, this means that the things by which we define events, the things which are the value of the variable z, are themselves events in virtue of being causes and effects, and so already presuppose a criterion of identity for that class of things.

But even if (DC) is not unacceptably circular, it is still unacceptable. Myles Brand gives the case of a particle which undergoes fission and then fusion [Brand,ibid., p.137]. Suppose that when fission occurs, one part spins to the left, the other to the right, and then they fuse without causally interacting with any other objects. *Ex hypothesi*, each spinning has the same causes and effects, and hence (DC) requires that they be identical in spite of the fact that the different spinnings involve different objects with distinct spatial location -- an untoward consequence indeed. Moreover, (DC) entails that there could be no possible world in which two distinct events have the same causes but no effects, no causes but the same effects, or that there be two distinct

events with no causes and effects at all.

However, there is another way to interpret Davidsonian events, recently endorsed by Davidson himself [1985a, p.175]. E.J. Lemmon had proposed that we view Davidsonian events as *space-time regions* or *temporal segments of objects* [Lemmon,1967]. This is, in fact, just the view Quine echoed long ago:

> Physical objects, conceived thus four-dimensionally in space-time, are not to be distinguished from events or, in the concrete sense of the term, processes. Each comprises simply the content, however heterogeneous, of some portion of space-time, however disconnected or gerrymandered [Quine,1960, p.171].

We might say that, on this view, events are understood to be "regions of occurrence," the event comprising everything within a specific region. Thus, according to the Lemmon/Quine/later Davidson view of events, or (LQD2), events x and y are identical when they inhabit the same spatio-temporal region or constitute the same temporal part of an object. More formally:

(LQD2) (x)(y) (x=y iff (z) (x is in z <-> y is in z)).

How this criterion differs from (DC), as far as circularity is concerned, is that the variable z does not range over events, and so does not presuppose a criterion of identity for that class of things. Instead, z is understood to either range over space-time regions or the objects whose temporal parts are the alleged events. [3]

Now contrast this view of events with the leading rival, the structural view of events developed by Jaegwon Kim [1966,1969,1973, and 1976]. Kim analyzes an event as the structured complex of an *object exemplifying a property at a time* . It is an object x having F at t rather than an object x at t pure and simple -- the temporal slice of x -- as on the previous view. In canonical notation we represent it as [x,F,t], where x is the "constitutive object," F the "constitutive property," and t the "constitutive time" of the event. This bracketed expression functions grammatically as a singular term, and is likened to a gerundive nominal, for example, "Jones' buttering the toast at midnight." But linguistic details aside, the bare metaphysical thesis of an event consisting of an object having a property at a time is extremely popular, having historical precedent in Leibniz [see Bennett,1988, p.92], and claiming a number of present day supporters [e.g., Goldman,1971; Lombard,1986; and Bennett,1988].

According to Kim, the criterion of identity for these structured events can be given along the following lines, where x and y now range over objects, F and G properties, and t and t' times:

(JK) $(x)(y)(F)(G)(t)(t')\ ([x,F,t,] = [y,G,t'] \text{ iff } x=y \ \&\ F=G \ \&\ t=t')$.

In other words, two events x and y are identical if and only if the objects, properties, and times which constitute those events are the same.

This concludes our presentation of the two leading theories of events. We have one theory, (LQD2), which represents what has been called a

"coarse-grained" scheme of individuation according to which events are *regions of occurrence* ; and we have another theory, (JK), which represents a "fine-grained" scheme of individuation according to which events are the *exemplifications of a property within such regions of occurrence* . The latter is more fine-grained than the former precisely because it focuses on the properties exemplified, and the differences between the properties, making for a distinction between events within all regions that exemplify more than one property (more on this in the next section). In any case, these theories will occupy us for the remainder of this chapter, each having different consequences for token physicalism.

Ramifications for Token Physicalism

The best way to see how these two views differ *vis-a-vis* token physicalism is to focus on the event types or properties exemplified by the objects which are involved in any given event. According to (LQD2), a particular event of type F will be identical to a particular event of type G just in case the property of being an F and the property of being a G are co-exemplified, that is, in the same spatio-temporal region or by the same temporal segment of an object. This is due to the fact that, again, an event is considered to be a spatio-temporal region or a temporal segment of an object, so that the exemplification of F and G in that region or by that object will suffice for the identity of the events falling under those types.

According to (JK), however, this co-exemplification is not enough to ensure an event identity. A particular event of type F will be identical to

a particular event of type G only if the properties of being an F and being a G are *themselves* identical, as the right hand side of the biconditional in (JK) makes clear -- assuming, of course, that the properties F and G are constitutive of the event, a point which we shall return to at length. Hence, on this view we need a *type identity* between F and G before we can identify the tokens of those types.

The ramifications for our study are chiefly these. To take a standard case, a particular C-fiber firing will be identical to a particular pain event, according to the more coarse-grained (LQD2), provided only that they inhabit the same spatio-temporal region. Of course, the precise location of many psychological events may be unclear. For example, we may want to locate the pain not merely at the C-fiber, but literally spread out to the source of injury (the pain "shoots down one's arm"). But no matter, for in such cases the psychophysical identity would still obtain, only now including a larger, more complex physical token. Hence, as long as mental events have a location at all, whether it be determined or not, then the individuation of events by (LQD2) will provide a way of vindicating token physicalism.

On the other hand, if the more fine-grained structural view is correct, then token physicalism will be extremely difficult to maintain. [4] A pain event and a C-fiber firing are identical only if the constitutive *properties* of being (or having) a pain and being (or having) a C-fiber firing are identical. But we have every reason to believe this is not the case. Quite generally, there is little hope in the idea that every mental property can be identified with some physical property. For, as argued in the previous chapters, the phenomenon of multiple realization makes it

unlikely that psychological properties are lawfully coextensive with physical properties; and hence they cannot be identical.

Of course, this incompatibility of the structural view with token physicalism assumes that *mental events have psychological properties as their constitutive properties* . For the identity conditions given by (JK) require the sameness of constitutive properties, the properties which the constitutive objects exemplify, and says nothing about the properties which the event structure itself may exemplify. That is to say, the aforementioned incompatibility will arise only because irreducible mental properties have been granted a constitutive status within structural events, thus preventing their identification with physical events. Indeed, it is precisely because of this role of constitutive properties that one can attempt to make the structural view consistent with token identities. Hence to that topic we now turn.

Modifying the Structural View: Nomic Properties

At this juncture we have two leading theories of events, (LQD2) and (JK), the former being consistent with token physicalism, while the latter appears not. I now want to consider a way of modifying the latter theory, the structural view of events, so as to avoid this consequence. In other words, is there a way the structural view can accommodate token identities ? Given the nature of structural events, this amounts to whether there are any good reasons for denying that mental properties can be constitutive of events.

Now, as far as I can tell, we would have good reason to deny that mental properties can be constitutive of structural events only if they

are somehow suspect in a way that physical properties are not. Put differently, we would have reason to deny that mental properties have a constitutive role within structural events only if psychological attribution is not taken seriously, or at least *as* seriously as the attribution of physical properties.

The most radical expression of this attitude, and what I believe to be the most implausible, is the eliminativist position [Feyerabend,1963a; and Churchland,1981]. For if we eliminate the category of the mental altogether, then there are no mental properties, and *a fortiori* no constitutive mental properties. However, it is my own working assumption that we should be realists about mental attribution, whatever our view of events and event identity may be [see, e.g., Horgan and Woodward,1985].

A less radical proposal, however, but one which would still make the structural view consistent with token physicalism, is a suggestion made by Terence Horgan on behalf of the physicalist [Horgan,1980 and 1981]. [5] The basic idea is to simply ban mental properties from playing a constitutive role within structural events on grounds that mental properties do not enter into any *strict scientific laws* , that is, laws which are "precise, explicit, and as exceptionless as possible" [Horgan,1980, p.667; the quote is taken from Davidson,1970, p.219]. Indeed, only the properties of physics have that privileged status. Thus, Horgan's proposal amounts to the claim that "every event consists in an individual's instantiating a natural kind of physics" [Horgan,1981, p.409].

The virtue of this proposal, over the eliminativist position, is that we do not deny the existence of mental events and properties. Rather,

mental events are *identified* with physical events because we have delegated mental properties to a lesser role with respect to the event structure [see Horgan,1980, pp.669-670]. In canonical notation, a mental event so modified is not [x,M,t], as before, but rather [x,P,t], which will itself exemplify M. But then the event [x,P,t] which exemplifies M is a physical event, ensuring the truth of token physicalism, even though M may be an irreducible, nonphysical property.

But why should the fact that mental properties do not enter into any strict laws prevent them from being constitutive of mental events ? What, exactly, is so important about strict laws ? And what is their connection with constitutive properties ?

Horgan attempts to supply an answer to these questions by suggesting three separate theses, from which the modified structural view will follow. They are:

(1) Mental properties do not enter into any strict scientific laws.

(2) Causation between events requires subsumption under strict law.

(3) All events are subsumed by law only in virtue of their constitutive properties.

Given (1) through (3), and assuming that mental events causally interact with other events, they jointly entail that no event has a mental property as its constitutive property. Thus, every mental event must have a physical property as its constitutive property (it must *be* a physical event), and this is just the modified structural view. So, in short, it is because constitutive properties are nomic properties, and

mental properties are not, that we are invited to accept Horgan's modification of the structural view.

Not surprisingly, this is reminiscent of Davidson's well known argument for "anomalous monism," for it is just that argument brought to bear on the structural view of events [cp. Davidson,1970 and 1973]. Hence the proposal is especially interesting in its own right, inasmuch as it supplies a direct argument for token physicalism. Unfortunately, I think neither Davidson's argument nor Horgan's proposal are well motivated, so let us look carefully at propositions (1) through (3).

As for (1), that mental properties do not enter into any strict laws, I take it that this should be granted on all sides. That is, if, like Davidson and Horgan, we mean by "strict law" one which is *exceptionless* , then there are no strict psychological laws [for more on "strict law," see McLaughlin,1985, pp.342- 348]. Indeed, any psychological law will have exceptions, perhaps built into the *ceteris paribus* clause, and for a variety of reasons: lack of attention, mental fatigue, and all manner of "systems failure" [see Fodor, 1974, pp.139-143].

I am also willing to grant, this time if only for the sake of argument, proposition (3) that structural events are subsumed by law only in virtue of their constitutive properties. Kim, in fact, made this suggestion in his own treatment of Humean causation [Kim,1973, pp.226 ff.]. And, putting aside for the moment questions about Humean versus other accounts of causation, I suspect that most proponents of the structural view would agree that constitutive properties should be our focus of attention inasmuch as the whole point, we may suppose, of invoking a structural event like [x,F,t] in matters of causal explanation is that we take

property F to be the causally relevant one in the circumstances, as opposed to some other property. On the other hand, I can imagine someone denying (3) by maintaining that the extrinsic properties of an event structure may also enter into causal relations. But I should like to put this question to one side, since it raises deep questions about the nature of structural events and how they causally interact. Hence, let us concentrate on proposition (2).

Proposition (2) says that causation between events requires subsumption under strict law. More informally, it says that strict laws have exclusive rights to causal power. Now the first problem I shall not press. But the second I take to be decisive. First, causation between events may not require subsumption under *any* law, strict or otherwise, which is just to say that Horgan's proposal is too closely tied to a Humean account of causation. We need not invoke the standard criticisms -- the cats, guns, and randomizing devices -- in order to make the point. The trouble is that a view is only as good as what it presupposes, and the argument for the modified structural view is presupposing a controversial view about causation.

Horgan is not unaware of this response. He notes, for example, that his proposal would lose all plausibility if we were to adopt a theory of causality along the lines Mackie has suggested [Horgan,1980, p.677]. For one event can be an "INUS" condition of another without any of their properties being related by law. And a similar point can be made, I might add, given the counterfactual approach to causation proposed by David Lewis [Lewis,1973]. For, on this view, causal relations are explicated in terms of counterfactuals whose truth conditions make no reference to

lawlike generalizations. Instead, the truth conditions are given in terms of an overall similarity between possible worlds, a similarity that is consistent with nomic divergence [see Lewis, ibid., p.184]. The point is simply that Lewis' theory does not support the claim that all causation is backed by lawful regularities.

But even if a broadly Humean view of causation is correct, still, why should we think that causation between events requires subsumption under *strict* law ? Oddly enough, neither Horgan nor Davidson address this issue in any direct way. [6] Horgan does say the following:

> I believe, with Davidson, that it is only within comprehensive, closed, scientific theories that we will find laws which are precise, explicit, and as exceptionless as possible. And laws with these characteristics must back singular causal statements, if we take seriously Kim's explication of Humean constant conjunction [Horgan, 1980, p.667].

But Horgan never tells us why, if we take seriously Kim's explication of Humean constant conjunction, it is only the strict laws which can back singular causal statements. Indeed, Kim's suggestion was just this: that structural events whose constitutive objects and times are spatio-temporally contiguous should be subsumed by virtue of their constitutive properties; and this has nothing to do with the *kind* of law which subsumes them, whether it be strict or exceptionless, deterministic or probabilistic, or anything of the sort.

Now, in the passage cited, Horgan also mentions the fact that strict laws are found within "comprehensive, closed scientific theories." So perhaps the idea is that there is a connection between causality and

systems which are comprehensive and closed in some suitably specified sense. Indeed, Davidson also refers to comprehensive, closed systems in his much discussed argument for the anomalism of the mental. [7] In his paper, "Mental Events," Davidson says:

> It is not plausible that mental concepts alone can provide such a framework [a vocabulary amenable to law] simply because the mental does not ... constitute a closed system. Too much happens to affect the mental that is not itself a systematic part of the mental. But if we combine this observation with the conclusion that no psychophysical statement is, or can be built into, a strict law, we have the Principle of Anomalism of the Mental: there are no strict laws at all on the basis of which we can predict and explain mental phenomena [Davidson,1970, p.224].

Here the inference is from the lack of strict psychophysical laws plus the fact that "the mental does not constitute a closed system," to the conclusion that there cannot be any strict psychological laws. Thus, suppose we know what a comprehensive and closed theory is supposed to be [see again McLaughlin,1985]. And suppose, further, that only basic physics is comprehensive and closed in the desired way. Still, Davidson's argument does not help us at all. We are already willing to grant his conclusion that psychological laws are not *strict* (again, in the sense of being exceptionless). That was proposition (1). The question is why this should be a problem; that is, why we should accept (2) which says that *causation* between events requires subsumption under strict law.

The fundamental issue, in other words, is whether causal powers are confined to the properties which enter into strict laws. If they are, then perhaps we should follow Davidson and look upon mental properties as having a lesser role with regard to causation and scientific explanation,

and follow Horgan and ban mental properties from being constitutive of events. But, again, why should we think that strict laws have exclusive rights to causal power ? For if there are psychological laws, albeit laws with exceptions, how does it follow that they are not *lawful enough to support causation between mental events* ? Grant everything Davidson says about closed systems and psychophysical laws. It has simply not been established that a theory (a) which does not constitute a closed system, and (b) whose predicates are incommensurable with respect to the natural kinds of physics, must be a theory without any serious laws, that is, serious enough to be invoked in causal explanation.

Indeed, assume the contrary. Assume that causality and the requisite laws are restricted to closed systems. The most damaging consequence is that this would rule out the possiblity of causation and law in biology and chemistry. For these theories also have predicates which are incommensurable with respect to basic physics [see, e.g., Hull,1974], and yet they do not constitute closed systems either, each being affected by intrusions from basic physics. Worse still, we cannot even say that *physics* is concerned exclusively with closed systems, as witnessed by the special attention paid to distubances of measurement within quantum mechanics [a point made by Suppes,1985, pp.184-186]. This is an especially interesting case, since it seems to be Davidson's point in reverse -- too much happens to the physical which is not itself a systematic part of the physical !

Finally, there are a legion of workers in psychology who would attest to the existence of serious psychological laws, statistical in nature, perhaps, but serious laws all the same, which are both empirically

correct and *no better or worse than the statistical laws of physics* [see Suppes, ibid., for examples from learning theory]. It does not seem plausible, then, to deny that psychological laws could support singular causal statements, at least not for the reasons given. Hence, we have yet to be given a reason to accept Horgan's proposal to ban psychological properties from being constitutive of mental events.

Parenthetically, my own conjecture about Davidson is that he supported a view which he believed to be detrimental to the causal/nomic status of mental properties because he inadvertently confused strict laws with the broader class of laws which are serious enough to be invoked in causal explanation. For it is only by glossing over this difference could one believe that the arguments which purport to show that there are no strict psychological laws are, on that account, arguments against the causal/nomic status of mental properties quite generally. Hence Davidson's unfortunate transition from talk about "strict" and "deterministic" laws [1974, p.230] to talk of "serious" laws which could connect reason with action [1974, p.233].

Modifying the Structural View: Normativity

There is, however, an interesting twist to the Davidsonian argument about mental properties which has recently gained some attention [see esp. Kim,1985]. According to this interpretation, there are indeed psychological laws of some kind. The reason is that mental attribution presupposes a viable psychological theory, and theories are individuated by virtue of their laws [Kim,1985, pp. 382-383]. What the mental does not have, however, are *predictive* and *explanatory* laws. Rather, the

laws which govern the mental are essentially *normative* in character, and it is this fact which serves to set psychological theory apart from the physical sciences. As Kim puts it:

> The view of psychology that emerges ... is one of a broad interpretive endeavor directed at human action, to understand its 'meaning' rather than search for law-based causal explanations that are readily convertible into predictions; psychology is portrayed as a hermeneutic inquiry rather than a predictive science [1985, p.383].

And in response to his critics, Davidson concurs in this emphasis upon normativity as the key to understanding his psychological anomalism:

> The basic reason the mental concepts connected with propositional attitudes cannot be incorporated in a system of exceptionless laws is the normative character of these mental concepts. Beliefs, intentions, and desires are identified by their objects, and these are identified by their logical and semantic properties. If attitudes can be identified at all, then, they must be found to be largely consistent with one another (because of their logical properties), and in tune with the real world (because of their semantic properties) [Davidson,1985b, p.245; see also p.249].

Ignoring Davidson's qualification about "exceptionless" laws, which, as we have seen from our previous discussion, cannot be the issue, the claim is now that consistency and overall reliability are normative constraints which must accompany psychological attribution, and that *these normative constraints are somehow inconsistent with the existence of predictive and explanatory psychological laws* . In other words, the problem, in Davidson's mind, results from a broad interpretive view of agents according to which we must find people to be generally

rational if we are, in fact, to attribute any beliefs to them whatsoever.

Now if this is true, then perhaps we would have good reason to adopt Horgan's original proposal after all. Namely, we should ban mental properties from playing a constitutive role within structural events, and thus preserve the desired token identities, on grounds that psychological properties have no predictive and explanatory use in a scientific theory. Hence, it becomes imperative to know precisely how such normative constraints are inconsistent with genuinely predictive and explanatory psychological laws. Unfortunately, Davidson has never been clear on this point. What is worse, most of his discussions are concerned exclusively with psychophysical laws, the aim being to refute claims about the theoretical reduction of psychology. So there is the additional burden of explaining how normativity at the psychological level is problematic not only for the existence of lawful connections between psychology and physical science, but also problematic for the existence of any lawful connections within psychological theory itself. Thus, I shall try to spell out the connection between these issues -- normativity, psychological anomalism, and psychophysical anomalism -- in a way that is at least in the spirit of Davidson's work.

First, it is clear that Davidson thinks the aforementioned normativity is inconsistent with the existence of psychophysical laws. He says that "there are no psychophysical laws because of the disparate commitments of the mental and the physical schemes," and again that "there cannot be any tight connections between the realms if each is to retain its allegiance to its proper source of evidence" [Davidson,1970, p.222]. Now the "disparate commitments" referred to which prevent such

inter-theoretic connections are precisely those which revolve around the normativity constraints. As Davidson says:

> Any effort at increasing the accuracy and power of a theory of behavior forces us to bring more and more of the whole system of the agent's beliefs and motives directly into account. But in inferring this system from the evidence, we necessarily impose conditions of coherence, rationality, and consistency. These conditions have no echo in physical theory, which is why we can look for no more than rough correlations between psychological and physical phenomena [Davidson,1974, p.231].

Thus, it is because an agent must be found to be largely rational in the way described, and because such normative constaints "have no echo in physical theory," that Davidson thinks there can be no genuine psychophysical laws. The reason, apparently, is that Davidson envisages the possibility of *conflicting demands upon psychological attribution* -- those which issue from the normative constraints within psychological theory, and those which would arise from the non-normative constraints within physical theory if there were to be psychophysical laws connecting the two realms.

Following Kim, we can say that the psychophysical laws in question would transmit purely non-normative conditions for attributing a mental property to the psychological theory, thus "preempting" or "seriously compromising" the role which normative constraints are thought to have in matters of the mind [see Kim,1985, pp. 375-381].

Precisely how the nature of this conflict is to be understood, or in what way psychophysical laws would preempt the role of any normative constraints, is a good question. But let us suppose, for now, that this is true. That is, let us suppose that the alleged conflict in normative and

non-normative attribution conditions would require that we deny the existence of psychophysical laws. How, then, would the argument against purely *psychological* laws proceed ? Can we derive psychological anomalism from psychophysical anomalism ?

One way to approach this question is simply to observe that, if there were no psychophysical laws, then it would be extremely difficult to see how there *could* be psychological laws. For example, on Fodor's explication of how the laws of a special science relate to those of a physical science, the existence of one-way conditional laws from physical theory to each special science type is actually presupposed [see Fodor,1974]. Indeed, the very idea that psychological properties are multiply realized *vis-a-vis* physical science presupposes lawfully sufficient (though not necessary) conditions within physical theory for each multiply realized psychological type. [8]

So it does seem as if psychological anomalism will follow, once we grant that there are no psychophysical laws. But should we ? More specifically, are Davidson and Kim correct in thinking that if psychophysical laws were to exist, then there would be conflicting demands upon psychological attribution ? In Kim's words, would such laws would transmit purely non-normative conditions for attributing a mental property to the psychological theory, thus "preempting" or "seriously compromising" the role of any normative constraints ?

I think not. It seems there could be *both* non-normative conditions for attributing mental properties along with the normative conditions which govern the mental. Brian McLaughlin, for example, has suggested that the normative conditions are logical in nature, arising from the concepts of

belief, desire, and action; whereas the non-normative conditions, in contrast, are purely empirical in nature, expressing a nomological connection with the mental properties. This being the case, then while we may acknowledge the *logical possibility* of a conflict in what the differing conditions may enjoin, the normative and the non-normative, we can say that such a conflict is *nomologically impossible* , as guaranteed by the existence of psychophysical laws. Put differently, to insist on the nomological possibility of a conflict is to deny the existence of psychophysical laws outright, and hence to beg the very question at issue [see McLaughlin,1985, p.358].

I should add, however, that there might be a different line of argument which Davidson presents in favor of psychological anomalism, though I am not at all confident that my interpretation is precisely what Davidson had in mind, or that he would accept my construal of the relevant passage. In his "Psychology as Philosophy," Davidson claims that there can be no serious laws connecting reasons and actions. He says:

> To see this, suppose we had the sufficient conditions. Then we could say: whenever a man has such-and-such beliefs and desires, and such-and-such further conditions are satisfied, he will act in such-and-such a way. There are no serious laws of this kind. By a serious law I mean more than a statistical generalization ... it must be a law that, while it may have provisos limiting its application, *allows us to determine in advance whether or not the conditions for application are satisfied* . It is an error to compare a truism like 'If a man wants to eat an acorn omelette, then he generally will if the opportunity exists and no other desire overrides' with a law that says how fast a body will fall in a vacuum. It is an error, because in the latter case, but not in the former, *we can tell in advance whether the condition holds* , and we know what allowance to make if it doesn't [1974, p.233, both italics mine].

I take it that by the phrase "determine in advance whether the conditions hold," Davidson means that we must be able to tell whether the background conditions hold independent of assuming the truth of the law in question. Thus, applied to the case at hand, the claim might be that, first, a psychological law must involve reference to other mental states in its background conditions; but second, and to bring in the familiar Davidsonian theme, given the contraints on rationality, consistency, and coherence, once we appeal to these other mental states in the background conditions, then more and more of the agent's beliefs and desires must be taken into account so that relationship between the original states expressed in the law is in fact already presupposed. In other words, by virtue of these holistic considerations, the truth of the psychological law is a foregone conclusion. Hence, unlike physical laws, there are *no theoretically independent means* , outside of assuming the truth of the psychological law in question, to determine whether its conditions for satisfaction hold.

Alexander Rosenberg seems to make a similar point in a paper entitled "Davidson's Unintended Attack on Psychology," only in this case Rosenberg speaks about the lack of theoretical independence for the "causal variables" of psychological theory rather than the lack of independence for its laws. As he illustrates with the case of phlogiston:

> One reason phlogiston theory came a cropper is that there were widely accepted ways of measuring the values of some of its variables, indirect ways, that were independent of phlogiston theory itself. After a certain point it became impossible to improve the theory's powers, or even reconcile it with experiment, without jettisoning these phlogiston-independent means of measuring its causal variables ...

> Intentional psychology does not have even the strengths of phlogiston theory, for its causal variables are not subject to any actual or possible independent measurement. This means not only that its explanations and predictions are incapable of improvement, but that its theory can shed no further light at all on the true singular causal statements about action and its determinants than "folk psychology" sheds on everyday affairs [Rosenberg,1985, p.404].

Either way, the point is much the same. Whether it be the assessment of its laws or the measurement of its variables, the normative contraints which are operative in all psychological attribution are viewed as preventing the kind of theoretical independence which we would expect from genuine science. And if this is true, then we have arrived at the point which Davidson and Kim have insisted upon in their discussions of normativity: mental properties fail to be predictive and explanatory in a way that physical properties are predictive and explanatory, in this case, because mental entities and the laws which govern them lack the theoretical independence afforded by physical science.

Now in order to assess this particular interpretation of Davidson, we need to know more about the kind of theoretical independence which science actually requires. As a start, and to speak more generally in terms of the attribution of properties, I think it is clear that scientifically respectable properties are such that the conditions for attributing them must be independent of the *particular* laws and theories which we might be interested to confirm. For example, "being an electron" is scientifically respectable because the attribution conditions are independent of, say, the particular theory of Lorentz on minimum

charges, as shown by the later experimental work of Millikan, or by the transition to a new theory of electrons like that of Bohr or Schrödinger [see Hacking,1983, pp. 83-84].

Notice, moreover, that the kind of theoretical independence illustrated by the case of electrons is perfectly consistent with the fact that the conditions for attributing the property are "theory laden." That is, we can still have a dependence on some theory or another. *Theoretical independence* , as I should like to say, does not require a complete *theoretical abstinence* .

To put the point in terms of observations used to confirm a particular theory, what this means is that genuine science requires only that an observation be independent of the particular theory whose confirmation is in question. If an observation is used to confirm T_1, then of course it cannot at the same time depend for its truth upon T_1. But it may well be dependent upon some other theory T_2. Indeed, this must be the case, at least for all those observations whose objects fall outside the "immediately given" in experience, if there is such a thing. So Millikan, when he measured the charge of an electron, did not presuppose the theory or hypothesis he was about to confirm. But he did presuppose *some* understanding of electrons, *some* background theory, for that is unavoidable.

Thus, the kind of theoretical independence which science requires does not entail a complete theoretical abstinence, either for the testing of its laws, the measurement of its causal variables, or more generally, for the attribution of its properties. But now we are in a position to cast doubt upon the foregoing arguments. Taking Rosenberg first, he said that:

> Intentional psychology does not have even the strengths of phlogiston theory, for its causal variables are not subject to any actual or practically possible independent measurement. This means not only that its explanations and predictions are incapable of improvement, but that its theory can shed no further light at all on the true singular causal statements about action ... than 'folk psychology' sheds on everyday affairs [ibid., p.404].

But here "intentional psychology" must be read as the *summation* of all scientific psychology, otherwise differing psychological theories, along with the distinct methodological procedures they suggest, would indeed provide the theoretical independence which results in the improvement of psychological predictions. Put in a different way, Rosenberg seems to have overlooked the fact that the desired theoretical independence could be supplied by *other psychological theories* . Indeed, a case parallel to our example with the electrons could be constructed for the attribution of belief, once connected to introspectionist theory, then refined and improved upon by the clinical methods of Freudian psychology, and then even more so by the experimental methods of present day cognitive science.

Hence, the mistake is to think that the entire practice of intentional attribution requires an independent grounding, with everything psychological being called into question at once (behaviorists fell into this trap, and so did Quine). We do not expect this kind of theoretical independence in physics, where the observational evidence for certain fundamental properties is unavoidably loaded with theory about what is fundamental; and so we should not expect our psychological theories to satisfy even more stringent requirements, as far as the objectivity and

independence of its postulates are concerned. The opposite tendency, in my view, would be equally mistaken -- the demand that physical properties have attribution conditions which are independent of all physical theory !

Davidson too, if I interpreted the relevant passage correctly, made a similar mistake. Davidson's concern was that a psychological law must refer to other mental states in its background conditions, which, given the contraints on rationality, consistency, and coherence, will include the rest of the agent's beliefs and desires so that the truth of the psychological law is already presupposed by virtue of these holistic considerations. Hence the condition for the law's satisfaction cannot be determined in advance, as Davidson put it, which is to say that the psychological law lacks the kind of theoretical independence required for scientific confirmation. But we can now see why this is wrong. For a psychological law L might belong to a theory T_1 whose confirmation is in question, while the independent conditions for attributing the mental properties involved in L might have been supplied by another theory T_2 -- precisely analogous to the situation Millikan faced.

The main difference between the two cases can only lie in the application of the normativity constraints. On the argument before us, reference to other mental states in the background conditions actually entails the truth of L *via* the holistic connections between the agent's mental states. But the normativity constraints need not be viewed in this manner. In the context of evaluating L of T_1, rationality demands, not that the agent already *have* the particular mental states involved in L (that could have been settled independently by T_2), but only that *if* the

agent has them, then they must be consistent with whatever other mental states the agent actually has.

The sum of the matter is that there are no especially compelling reasons to believe in psychological anomalism, or to believe that mental properties fail to be predictive and explanatory like good scientific properties. Hence, I conclude that there are no good reasons for adopting Horgan's proposal to ban mental properties from being constitutive of structural events. [9] Indeed, I think there are a number of positive reasons for believing that psychological properties are constitutive of mental events, of which the following should suffice.

Constitutive Psychological Properties

First I think there is a strong *prima facie* reason for thinking that mental events should have psychological properties as their constitutive properties. After all, given that an event is an object exemplifying a property at a time, as the structural view holds, then for each property exemplified by an object there would seem to be a corresponding event with that property as a constituent. It follows that if an object exemplifies a mental property at all, there will be a mental event having that property as its constitutive property.

A more telling argument, however, can be made for the view that a psychological property should be constitutive of a given mental event precisely because (a) the psychological property is essential to its identity, and (b) its microphysical properties are inessential. Hence, assuming that the essential properties are the best candidates for being constitutive of structural events, then psychological properties will win

by default, being constitutive of mental events. The argument will be adapted from what Richard Boyd calls the "transworld compositional plasticity" of token events [Boyd,1980, pp.99-101].

First we begin with the claim that what is essential to the identity of a particular mental event is indeed something psychological. One psychological feature, for example, is the *functional role* an event plays in the cognitive system of the subject. This claim seems well supported by present-day psychological theory, and cognitive science generally, since mental states, or at least the information bearing ones, are typically defined by their functional roles. [10]

Now for the second part of the argument, the claim that the physical properties of a mental event are inessential, Boyd argues for a more general truth that the microphysical composition of all macro events are inessential to their identity. Consider a common object like a car. It can survive the replacement of at least some of its parts, for example, the generator. So it would be the same car even if, counterfactually, it had a different generator. Moreover, we assume that it would be the same car even if, counterfactually, this new generator were made of a different substance with a different microphysical constitution.

But now consider an event involving the same car, for example, its spinning off the road. By the same reasoning it should remain the very same event even if, counterfactually, the car had the different generator made from an entirely different substance. But this counterfactual case is one in which the *molecular constitution of the event has changed*, if ever so slightly, and the point is perfectly general. We can imagine additional changes, so that there is no particular microphysical

constitution which is essential to the event.

Boyd concludes that the same thing must surely hold for the mental, so that their microphysical composition is likewise inessential. Thus, like mental types, mental tokens can have different transworld realizations (though only mental types can have different actual world realizations). And this being the case, it is the mental property and no microphysical property which is essential to the identity of a given mental event. This should, I think, provide a compelling reason to think that, on the structural view of events, psychological properties will be constitutive of mental events.

Indeed, this view about the transworld identity of mental events is extremely interesting in its own right, supplying an additional argument against token identities quite apart from our present concerns over structural versus nonstructural views of events. For if a mental event does have the appropriate transworld compositional plasticity, then it will have counterfactual properties which distinguish it from any microphysical event. Specifically, the mental event will have different counterfactual properties because any microphysical events associated with it in the actual world will not occur in those other worlds where the physical constitution of the mental event has changed.

Finally, one more argument in support of constitutive psychological properties, this one derived from linguistic considerations. There has been much ink spilled over the question as to when two descriptions refer to the same event. Kim's official answer seems to be that two descriptions pick out the same event if and only if they lead to the same explanations; or, alternatively, that a difference in the explanatory

power of the descriptions is both necessary and sufficient to show a difference in constitutive properties [see esp. Kim,1976; and cp. Lombard,1986, pp.50-55]. The trouble, of course, is that even if a difference in explanation suffices to establish a difference in constitutive properties, this by itself does not tell us what the different constitutive properties *are* (perhaps, in the case of psychological descriptions, they are two different *microphysical* properties). Hence we still lack a general procedure for determining what constitutive properties are expressed by psychological descriptions.

But here a suggestion by Lawrence Lombard can be brought to bear on our topic [Lombard,1979 and 1986]. According to Lombard, a description of a structural event is a canonical description, like Kim's "[x,F,t]," only when the "F" is an atomic event verb as determined by some scientific theory. In other words, it is only those descriptions relativized to a scientific theory which pick out the constitutive properties of events [see 1986, pp.166-177]. So the descriptions which express the constitutive properties of events are the *basic predicates of a science* , and the constitutive properties are the *basic properties attributed by that science* .

If this is correct, then all we need affirm is that cognitive science is a genuine science. It follows that the basic predicates of psychological theory -- talk of sensations, beliefs, strength of desire, scripts, frames, and all the descriptions of functional architecture -- will pick out corresponding mental properties which are constitutive of the events being described. That is to say, and assuming that the psychological predicates are not reducible to physical theory, basic psychological

predicates will pick out psychological properties which are constitutive of mental events.

Of course, one might object that Lombard's suggestion is much too strong, at least as he has formulated it. To be sure, it is an improvement over Horgan's proposal which confines constitutive properties to microphysics. All the same, why accept *any* restriction to scientific properties ? For example, macro properties like "being a bridge" are not found within the various branches of science, and neither are they reducible to them. Nevertheless, we might think that there are events involving things like bridges, even having the irreducible macro properties as their constitutive properties.

Now if we do think this, then no harm will follow. We can simply take Lombard's suggestion as a *sufficient* condition for determining constitutive properties, though not a necessary one. The result is still the same: for every basic predicate of a scientific theory there will be a corresponding constitutive property for the appropriate events. And since there are basic predicates of a scientific psychology, then there are constitutive psychological properties for mental events.

Thus I conclude that psychological properties are constitutive of mental events. As a consequence, the criterion of identity for structural events will require that psychophysical tokens to be of the same type, which is just to say that if token physicalism is true, then mental and physical properties must be identical. Hence, given that they are not identical, as argued in the previous chapters, then the structural view of events will not allow for token identities.

In summary, then, we have two leading theories of events: one which views events as spatio-temporal regions or temporal slices of an object, the other which views events as structured complexes of an object exemplifying a property at a time. Our present investigation has shown, with a fair degree of plausibility, I think, that the two theories do indeed have different consequences for token physicalism, one favorable and the other not. My strategy, to be carried out in the next chapter, is to rule out that theory which looks favorably upon token identities.

Notes to Chapter III.

[1] I do not mean to sound cavalier about the choice of a linguistic theory. Some may indeed carry different metaphysical commitments. E.g., Horgan analyzes causal statements in terms of a nontruth-functional connective which relates whole sentences, rather than Davidson's way of using a predicate to relate singular terms [Davidson,1967b; Horgan,1978]. By doing this, Horgan claims that we can eliminate the need for events. However, I am not convinced. Granted, causes and effects will no longer be denoted by singular terms. But they will presumably correspond to the complete sentences, and tradition has it that *facts* or *states of affairs* play this role. Therefore we need to know more about these entities before we exclude events (i.e., perhaps events *are* facts or states of affairs, or a subclass of them, the concrete ones). Thus, there is a very large and difficult metaphysical issue to be resolved about the nature of events vs. facts before we could know what the ontological consequences are which accompany the different semantic theories. Even Bennett, who goes to great lengths to distinguish fact and event talk [1988, chapters I and II], ultimately concludes that events are a kind of fact which are described in a less specific way [ibid., esp. pp.128-130]. Moreover, it is worth noting that on some linguistic theories the facts or states of affairs come out looking much like structural events [e.g., see Clark,1970, p.331].

[2] I agree with Thalberg on the need for events. However, I do not agree that Kim's structural view is *reductionist* in spirit, eliminating their category altogether as Thalberg claims. It *can* be reductionist if we interpret the event structure as an ordered triple, i.e., as a *set* consisting of an object, property, and time. But we need not interpret Kim's theory in this way. Indeed, we had better not, since the aforementioned set could exist even when the object does not exemplify the property in question [see Kim,1976, p.161; and esp. Bennett, 1988, p.91].

[3] All three philosophers mentioned seem indifferent between the two possible readings, that is, using space-time regions versus temporal segments of objects, although Quine and Davidson seem to prefer the first manner of speaking [e.g.,1985, pp.167 and 175 respectively]. For the time being, however, we can ignore this complication. It shall be taken up in chapter 4.

4 Token physicalism will be exccedingly difficult to maintain unless certain restrictions are placed on the structural view. One restriction, which I critically discuss in the next section, allows only the properties of *basic physics* to be the constitutive properties of events. Another restriction, suggested to me by William Taschek, allows only the *spatio-temporal* properties to be constitutive of events. This latter suggestion is especially interesting because it expresses the doctrine of events as temporal slices of objects in terms of Kim's structural view. In any case, the arguments made against the temporal slice view in chapter 4 will apply equaly well against this modified structural view.

5 It should be emphasized that this is only a suggestion on behalf of the token physicalist. Horgan's own view is that there are no events [Horgan,1978]. But he does not deny that objects exemplify properties, even mental ones. And, in fact, even if there are events, Horgan would reject token physicalism on other grounds [Horgan and Tye,1985].

6 In "Mental Events," Davidson claims that: "events related as cause and effect fall under strict deterministic laws" [1970, p.208]. But he goes on to add that the principle will be treated as an *assumption* , and then refers us to another paper in which the notion of causation is explicated [1970, p.208 and fn.3]. However, that paper is "Causal Relations" [1967b], and i find nothing there to support the idea that causality must be confined to strict laws. Notice, e.g., that it cannot follow from the fact that the nomological character of causality invites a nomological deductive account of causal explanation -- the idea being that only strict laws allow us to *deduce* the explanandum event from the laws and the initial conditions. For we could just as well use *probabilistic* laws to account for whatever exceptions obtain with psychological generalizations, and adjust our inferences accordingly.

7 Horgan seems to accept Davidson's argument, for he makes a similar inference from the lack of psychophysical laws to the "nomological incongruity" of everyday psychological properties. The only difference I find is that Horgan adds the premise that such everyday properties "were never intended to figure in the laws of a precise theory" [1980, p.668]. But Horgan should not press this point, at least not the most recent Horgan who defends the use of folk psychological concepts in cognitive science [Horgan and Woodward,1985]. Indeed, the same thing could be said about a number of terms in the physical sciences, specifically, those taken over from folk biology and folk physics. Hence, nothing which Horgan has said

will forbid the employment of mentalistic predicates in a precise theory, and much of cognitive science has attempted to do just that.

[8] See again my chapter 1., p.5.

[9] Horgan does give other reasons to support his modification of the structural view, namely, its "theoretical simplicity" [Horgan,1980, p.672], and its avoidance of "higher-level events which would either duplicate the causal role of events or else would dangle from the causal nexus as mere epiphenomena" [Horgan,1981, p. 411]. In other words, the modification provides for a simpler ontology, since token identities eliminate a separate realm of mental events; and, by eliminating them, it avoids the problem of explaining how these separate and irreducible events have a causal role to play alongside that of the basic physical events. As for the second point, Kim has provided an alternative picture of how, on a dualistic view of events, the mental causally interacts in virtue of its supervenience upon the physical [Kim, 1979,1984]. More important, I think, is the first point about theoretical simplicity. It should be observed, however, that by accepting a dualism of properties (as Horgan does) *we still have a significant theoretical complexity* -- distinct levels of properties and explanatory schemes. Finally, the appeal to Occam's razor is decisive only when the competing theories cannot be decided on other grounds, i.e., only when *all things are otherwise equal*. But this is not the case here, as I hope to show in the next section.

[10] See again my chap.1., fn. 5, for an important proviso about psychofunctional identity.

CHAPTER IV.

THE PROBLEM WITH TOKEN IDENTITIES

In the last chapter we were left with two leading theories of events: a coarse-grained view endorsed by Lemmon, Quine, and the later Davidson, according to which events are spatio-temporal regions of occurrence or temporal slices of objects; and a fine-grained view endorsed by Kim, Goldman, Lombard, and Bennett, according to which events are structured complexes of objects exemplifying a property at a time. Now, as I see it, the coarse-grained view is simply mistaken. It cannot plausibly be taken to supply an analysis of mental events which are needed for the cognitive sciences, nor, I think, for the events of any other explanatory enterprise. [1] The structural view, on the other hand, seems to fit our talk about events in the best way possible, and so, barring any serious difficulties, it is the preferred theory of mental events. This being so, and given the incompatibility of the structural view with token identities, as argued in the previous chapter, then it appears we must reject token physicalism.

As a prefatory remark, it should be noted that this general line of argument is not entirely new. I have pieced it together from a number of sources, beginning with some early remarks of Kim which concern the difficulty of maintaining token identities on his theory of events

[Kim,1966], coupled with my rejection of those proposals which seek to modify the structural view in order to accomodate such identities [as in Horgan,1980 and 1981], and combined with certain problems to be discussed shortly which face the coarse-grained view of events.

Some philosophers have, in fact, recognized that the plausibility of token physicalism rests upon our choice of the coarse-grained view [e.g., Feldman,1980]. But they have, in all appearances, underestimated those difficulties which I take to count decisively against the view. It is not, I think, a matter of choosing one's weapon, but of finding one which is reliable enough to withstand a charge.

(LQD2) and Spatially-Overlapping Events

To begin our criticism, I want to distinguish two readings of the coarse-grained view of events. [2] According to this view, events x and y are identical when they inhabit the same spatio-temporal region or constitute the same temporal part of an object. More formally, the criterion is:

(LQD2) (x)(y) (x=y iff (z) (x is in z <=> y is in z)).

On the first reading we let z range over *spatio-temporal positions* , and on the second we let z range over *temporal segments of objects* . The two are materially equivalent on the assumption that only one object or substance can inhabit a particular spatio-temporal region. But I think that assumption is mistaken, as we shall see. In any case, let us interpret (LQD2) in the first way as individuating by spatio-temporal

position.

As a start, it should be observed that there is nothing wrong with the idea that distinct things can spatio-temporally overlap. David Wiggins, most notably, has argued for this in the case of *objects* and their *substance* or *matter* [Wiggins,1967 and 1968; also Shoemaker, 1970]. To take a standard case, a statue spatially overlaps the gold which constitutes the statue throughout its career, yet the two remain distinct because of their different "life histories," which is just to say that the gold existed before the statue and will continue to exist after the statue has perished. But this means that they have different properties, and so by Leibniz Law they cannot be identical. Moreover, even if their life histories happen to overlap, there is still a difference in their counterfactual properties. For example, the gold *could have existed* apart from the statue, if the gold had not been so formed, and the statue *could have existed* apart from the gold, had its pieces been slowly replaced over time.

Furthermore, this sharing of spatio-temporal position can occur not only between objects and their matter, but also between *ordinary objects*. An example from Burge is the case of a rope and a hammock made from the rope [Burge,1975, p.462]. The two occupy the same position during the time at which the hammock exists, and yet they are distinct because of their different histories and various other properties, modal and otherwise.

Finally, this overlapping can also occur between ordinary objects of the *same kind* (though perhaps not throughout the entire career of both objects). For example, Leibniz gives the case of two shadows which

cross each other's path [cited in Sanford,1970, p.75]. Imagine two separate light sources, for example, which happen to cast shadows in the same place at a particular time. There is good reason to think there are two shadows at the place where they overlap, or more precisely two temporal parts from each of the two shadows, since the very same conditions hold for the existence of the temporal parts in the place of overlapping as there is for the existence of other temporal parts in other places. These conditions include the fact that there are two distinct objects which cast the shadows, the different causal explanations for their existence in those places, and the like, all of which remain operative when the shadows overlap.

Or consider a different example, this one involving more concrete objects. Imagine two waves which, starting from different directions, cross each other's path [Haugeland,1982, p.100]. At the point where the waves intersect they share the same position, and so there are two waves (or two wave parts, one for each wave) at that position. Again, we have the very same reason to believe that there are two temporal parts at the point of intersection, one from each wave, as we do for believing that there are wave parts at other places. In this case, there are the different origins of the waves, their distinct velocities, their continuation in different directions, and the like, all of which indicate a continuous succession of wave parts including the place where they overlap.

Since I shall make much of this phenomenon in due course, let me pause to reinforce this point. It might be suggested, in response, that we should describe the above cases somewhat differently. Namely, we

should describe them as situations according to which there is only *one* temporal part *shared equally* by the two overlapping objects. [3] It seems to me, however, that this way of viewing the matter simply ignores the facts I have already stressed, namely, that we have the same reason for holding that there are two temporal parts in such cases, one for each object, as we do for saying that there are temporal parts of the objects at any other place.

Moreover, given that it is not metaphysically necessary that the objects overlap at the time in which they do, it seems clear that the temporal parts of each object will have different counterfactual properties, thus preventing their identification. For example, where O_1 and O_2 are the objects which spatially overlap at some time T, the temporal part "O_1 at T" has the property that it *could have existed apart from* O_2, unlike the temporal part of O_2; and the temporal part "O_2 at T" has the property that it *could have existed apart from* O_1, unlike the temporal part of O_1. Consequently, I take it that our best description of the situation is that there are indeed two overlapping temporal parts.

Now the latter cases I have mentioned actually parallel the situation with events quite nicely. For I want to argue that just as the rope and the hammock or the two waves each spatially overlap by sharing the same matter (some rope fiber in the one case, H_2O in the other), so also can two events spatially overlap by sharing the same underlying object. Quite generally, then, objects can overlap on the same matter, and events can overlap on the same object.

Hence, turning from objects to events, notice that if objects can spatially overlap at certain times, then so also will the events or states

of those objects which occur at those times. This is true on the assumption that an event is an object exemplifying a property at a time, since at the time when two objects overlap there will be two events in virtue of the *different constitutive objects* ; and it is even true on the assumption that an event is a temporal segment of an object, since at the time when two objects overlap there will be *two temporal segments* at that position, one for each object.

What cannot be true, however, is (LQD2) when it is interpreted in the first way as individuating by spatio-temporal regions. For assuming that objects can overlap in the way described, and that the events or states of the one object are not identical to the events or states of any other object, then at the time when the objects overlap there will be those different events for each object occurring in the same region, contrary to (LQD2). Therefore (LQD2) must be interpreted in the second way according to which events are temporal segments of objects. Henceforth we shall understand it in this second way.

Now in order to refute the view that events are temporal segments of objects we should need clear cases in which, unlike the previous ones, (a) there need be only one object under consideration which is the subject of an event, that is, no overlapping objects which could allow for distinct temporal segments occurring at the same time, and yet (b) this one object's temporal parts are nevertheless inadequate for the individuation of the events involving that object. In simplest terms, we need to find *a single temporal part of a single object which sustains more than one event* . Such cases are not hard to come by, and, in fact, Davidson himself discussed such a case when he first spoke of Lemmon's

proposal:

> Doubt comes easily in the case of events, for it seems natural to say that two changes can come over the whole of a substance at the same time. For example, if a metal ball becomes warmer during a certain minute, and during the same minute rotates through 35 degrees, must we say they are the same event ? It would seem not; but there may be arguments the other way [Davidson,1969, p.178].

Here we have just one object, the ball, which is the subject of two separate events or states at the same time, its becoming warm and its rotating. So the one temporal segment of the ball appears to sustain two different events, and this is contrary to (LQD2) which requires the identity of all those events which are confined to the same temporal slice of an object.

Now, as a matter of fact, Davidson did suggest an argument the other way, claiming that the spinning and the heating of the ball might be the same in virtue of being identified with the same molecular motion. But this is certainly not compelling. First, it seems to beg the question, for unless we already presuppose the coarse-grained view of events, there is no reason to identify either the spinning or the heating with the molecular motion. Second, even if the spinning and the heating are identical to some molecular motion (instead of being supervenient, for example), it does not follow that they are identical to the *same* molecular motion [see Brand,1976, p.145; and Bennett,1988, p.113].

In any case, the example Davidson discusses is actually far too generous for a counterexample. What I mean is that in the case where the ball spins and becomes hot, we do happen to think there is a connection

between the two events, and this may, perhaps, motivate us to identify them at some deeper molecular level. Nevertheless, it follows from (LQD2) that *any* events or states of an object occurring at the same place and time will be identical. Not only will the spinning be identified with the heating, but so also will the following be identical:

(i) the ball spinning at t,
(ii) the ball being orange at t,
(iii) the ball being round at t.

And the same holds for another mixed trio [cp. Davidson,1980, p.125]:

(iv) my swimming the Channel at t,
(v) my shivering at t,
(vi) my turning blue at t.

None of these, I think, can plausibly be said to constitute identical events in spite the fact that they occur in the same object and at the same time. [4] Of course, in the latter cases it may be unclear whether the subject of the event is a *person* or a *body* , and this, arising from the fact that swimming is an intentional act whereas shivering and turning blue are not. And this is to say that there may be more than one temporal part at work here, one for the person and one for the body. Yet other examples could be multiplied at will, and I am fairly confident that one could not come up with enough objects and temporal parts to account for every one of the apparently distinct but concurrent events.

Moreover, no such problem affects ordinary object attribution, like cases (i) through (iii), since it is clear that one and the same ball is the subject of each of the aforementioned events (if you have doubts, suppose that the ball is colored all the way through so that the *entire* ball is orange, the *entire* ball spins, and so on). Thus, since we cannot believe that the ball's spinning, its having an orange color, and its being round in shape, are identical events; or indeed, that any object's movement, the having of its color, and the having of its shape are identical events, then any theory which entails such a thing should be rejected.

Physical Shareability of the Mental

A similar trouble will arise over spatio-temporal location, only this one having to do specifically with mental events. First let us suppose, with Davidson, that a mental event is located at the person, which is to say that it is the person who is the subject of the mental event [Davidson,1969, p.176]. Now one thing we should certainly want to say about mental events is that a person can have *several at once* . My thinking of Vienna while I swim the Channel can occur at the very same time that I desire to finish the swim or hope for a St. Pauli's Girl to reward my efforts. But given Davidson's way of locating mental events at the person, and given (LQD2), we obtain the result that *my thinking of Vienna is identical to my desiring to cross the Channel and my hoping for a St. Pauli's Girl.* Indeed, the point is perfectly general -- *all* mental events co-exemplified by a person are identical !

This, I take it, would be disastrous for cognitive science. The problem, of course, is that a temporal slice of a person can sustain an indefinite number of mental events. Thus we cannot individuate a person's mental events in terms of its temporal parts. Something must give, and I suppose that once Davidson traces out the consequences of (LQD2) he will no longer locate mental events at the person. The alternative, however, given (LQD2), is to believe that each concurrent mental event has its own unique location; and this is presumably to say that the subject of each mental event is a distinct neurophysical object, this C-fiber for that pain, this neurological structure for that thought, and so on. But this alternative, in my view, is not open. As we shall see, it presents an implausible picture of the mind, especially given the computational approach to cognition which has been adopted by nearly all contemporary psychological theory.

Thus, I have argued that a mental event cannot be a temporal segment of an object when that object is a person. I will now argue that a mental event cannot be a temporal segment of an object when that object is some neurophysical structure. Assuming this exhausts all plausible alternatives, then we will have disposed of our principle (LQD2) as applied to mental events.

Why should we think mental events cannot be individuated in terms of the temporal parts of neurophysical objects ? Because the following seems true, call it the *Physical Shareability of the Mental* :

More than one mental event or state can have the same physical basis at the same time .

Remember that we are still talking about concrete events, that is, mental tokens, not types. This contrasts, then, with the phenomenon of multiple realization. For whereas multiple realization concerns a single mental *type* being realized by different underlying physical tokens on different occasions, physical shareability concerns more than one mental *token* being subserved by the same physical token at the same time.

In any case, if this is true, if, indeed, a single brain structure can sustain more than one mental event at the same time, then we cannot individuate mental events by counting them as temporal segments of objects, and hence (LQD2) must be rejected.

Unfortunately, I do not have a completely decisive argument for the Physical Shareability of the Mental. But I think it is plausible, and for a number of reasons. First an argument derived from Dennett, and developed in more detail by John Haugeland [Dennett,1976, p.107; and Haugeland,1982, pp.101-102]. Suppose we have a standard chess-playing computer. Or better yet, suppose I am playing chess, and that my thinking processes are roughly analogous to that of a computer. After observing my behavior, one can justly ascribe to me the desire to get my queen out early, the belief that I prefer a wide open game, that I like aggressive play, and so on, which is to say that one can can attribute to me the following mental states:

(vii) my wanting to get my queen out early at t,

(viii) my preferring a wide open game at t,

(ix) my contemplating aggressive moves at t.

Now the point is that the existence of such mental events or states does not depend upon my having distinct neurophysical structures for each one. As Dennett put it in the case of the computer:

> But for all the many levels of explicit representation to be found in that program, nowhere is anything roughly synonymous with "I should get my queen out early" explicitly tokened. The level of analysis ... describes features of the program that are, in an entirely innocent way, emergent properties of the computational processes that have "engineering reality" [ibid., p.107].

So, like the expert system just described, perhaps the only thing we find explicitly represented in my mind are certain heuristics for evaluating chess positions. But, to press the computer analogy, there need be nothing in my data structures which corresponds to any one of the mental states mentioned in (vii) through (ix) to the exclusion of the others. On the contrary, they emerge together as "net attitudes," in Haugeland's words, once the explicitly represented heuristics are fixed.

Parenthetically, this view has also been taken up by members of the artificial intelligentsia. Winograd, for example, wants to move away from those models which have a distinct mechanism or data structure for every intentional state represented in the system. Much like Dennett he says:

> If I say of a program, "It has the goal of minimizing the number of jobs on the waiting queue," there is unlikely to be a "goal structure" somewhere in memory or a "problem solving" mechanism that uses strategies to achieve specified goals. There may be dozens or even hundreds of places throughout the code where specific actions are taken, the net effect of which is being described [Winograd,1981, p.250].

Indeed, once we take seriously the computational approach in psychology, then the identification of our propositional attitudes with the net attitudes seems quite natural. That is to say, for at least many of our common propositional attitudes like my wanting to get my queen out early, my preferring a wide open game, and so on, the computer analogy may force us to treat them as net attitudes which emerge together from a lower level of explicit representation. And, if this is true, then many of our common beliefs and desires will jointly share in the same explicit representations, and hence share in the same physical basis which happens to underlie those representations. In a word, the Physical Shareability of the Mental will be a direct consequence of computational psychology. Hence, let us turn to this important topic.

Inexplicit Content and Giobalism Considered

Robert Cummins has recently described the Dennett-Haugeland case as an instance of a more general phenomenon within computational psychology which he calls "inexplicit content" [Cummins,1986 and 1989]. Seen from this perspective, the aforementioned net attitudes are the inexplicit content of a cognitive system which differs from the content explicitly represented by the system's data structures and which arises out of the system in various ways.

Consider, for example, what Cummins refers to as *content implicit in the state of control* :

> A word processor's search routine tries to match the character currently being read against the second character of the target only if the character read last matched the first character of the target. If it is now trying to match the second character, the

> current state of control carries the information that the first character matched the last character read; however, the system creates no data structure with this content. Nowhere is that content explicitly represented [Cummins,1989, p.16].

Thus, to use a simple illustration, if the target of the computer's search is the word "Fred," and the last character read was the letter "F," then the search routine carries this information when it tries to match the second letter "r." But there is no data structure within the system (no explicit representation) which can be interpreted to mean: "the last character read is the letter 'F,' which matches the first letter of the target word." Rather, it is the overall behavior of the machine which justifies the attribution of that particular content to the system. In roughest possible terms, the content has more to do with the system's control box, not the belief box.

Or consider a different case, not so closely aligned with computers, what Cummins refers to as *content implicit in the domain* :

> I give you instructions for getting to my house from yours, all in such terms as "go left after three intersections' and "turn right at the first stop sign after the barn" ... Now, if you (or anything else) execute this program, you will get to my house. In the process, you never create a representation of the form "Cummins lives at location L"; yet, given the terrain, a system executing this program does "know where Cummins lives" [ibid., p.17].

This particular case is more like the net attitudes described in the preceding section. For the mental content expressed by the proposition "I know where Cummins lives" appears to supervene on the system in virtue of what is explicitly represented (in addition, perhaps, to the

environmental context which enables it to display the knowledge in question), but without that content being identifiable with any explicit representation. And this naturally suggests the following picture, which Cummins labels with the term "Globalism"; namely, that *the propositional attitudes are global states of a cognitive system* .

On this view, as Cummins puts it, "a belief stands to the [computational theory] supported psychology as a "point of view" stands in relation to an editorial: There may be no particular bit of the editorial that expresses the point of view; the whole editorial does it" [1989, p.143]. The only alternative, however, at least within the general framework of computational psychology, is to identify each propositional attitude with its own separate data structure (as opposed to a global mapping of the attitudes onto the data structures), where these data structures are now the explicit representations of whatever propositions are believed or desired. Hence, call the alternative which identifies particular mental contents in a one-to-one fashion with distinct data structures "Exclusive Localism." [5]

Now it seems to me that Exclusive Localism is mistaken, and Globalism clearly preferred. First, a point against Exclusive Localism, there simply are not enough data structures for every propositional attitude which can be attributed to a cognitive system, as the examples of the net attitudes will show. Second, and also a point against Localism, there are many data structures which are not even plausible candidates for the propositional attitudes, for example, the 2 1/2 dimensional sketches employed within computer vision or the phonological representations invoked by speech-recognition theories [Cummins,1989,

p.143]. And third, unlike Exclusive Localism, Globalism can explain the general looseness of fit and relativity of belief attribution emphasized by Dennett [1978] and Stephen Stitch [1983].

For, as Cummins observes, when we attribute a point of view to an editorial, some cases "will leave room for rational observers to disagree, to be uncertain, and to fall back on pragmatic considerations and rough estimates" [ibid., p.144]. Therefore, given the Globalist picture, attributing propositional attitudes will be very much the same, involving the same degree of uncertainty and relativity which is illustrated by editorials and the points of view they express.

So, to summarize thus far, we have looked at what I take to be a plausible case for the net attitudes, these being certain mental events or states which share a physical basis in virtue of the fact that they emerge together once an explicit level of representation has been fixed. And we have now seen that the net attitudes appear to be a direct result of the computational approach to cognition, since they form part of a wider class of psychological states which Cummins refers to as the class of inexplicit contents. Moreover, we have seen that such contents are best viewed as global features of a cognitive system.

This latter fact is of paramount importance. For if Globalism is correct, then the Physical Shareability of the Mental will be confirmed in a surprising way. Mental phenomena can share the same physical basis, in this case by globally sharing *all* of the physics which underlies the cognitive system ! Consequently, and this is the main point we are concerned to establish in this chapter, the individuation of events by (LQD2) must be rejected since it requires a distinct physical object --

no physical sharing -- for every contentful mental state which may exist.

Dispositional or Constant Belief States

One might be skeptical, perhaps, if the case for the Physical Shareability of the Mental rested entirely upon the net attitudes and the inexplicit contents just discussed. For we might think that explicit representation for each intentional state is the core of mental attribution. On the other hand, the idea of a core of explicit representation upon which certain net attitudes supervene strikes me as an attractive picture, and a quite natural one if our mental processes involve anything like the cognitive architecture described in the preceding sections.

All the same, we can appeal to more than the net attitudes. This view about physical shareability is especially plausible with respect to ordinary dispositional beliefs, or constant belief states, and these are the stock-in-trade of mental attribution [see Armstrong,1973, chap.2]. Consider, for example, beliefs about large mathematical sums. Clearly I do not have all the numbers explicitly represented, they are much too many and too large. Rather, I only have a certain amount represented, along with certain basic mathematical functions, from which I am disposed to generate the larger sums. Moreover, the same explicitly represented numbers and functions can be used to generate different sums. For instance, the explicit representation of both the number 10 and the multiplication relation can be used to generate different sums in the following problems: "10x10x10 = 1000" as opposed to "10x10x10x10 =

10000." But this means that our *corresponding dispositional beliefs about these sums have the same basis* , that is, the same numbers and functions, and hence the same *neurophysical basis* .

Or, not to belabor the point, think of all the different dispositional beliefs or constant belief states which are constituted by the same basic stock of concepts with the same recursively specified syntactic rules. The same representations will be involved, even though, dispositionally, one has them built into larger and more complex beliefs. So, in short, not only the net attitudes but also our dispositional beliefs and desires are such that they can be grounded in the same underlying physical facts.

This phenomenon of mental events sharing a neurophysical basis should not be surprising. There is no reason to suppose that the ontology of our psychological theories should mesh in a precise way with the ontology of our physical theories. Indeed, it would be miraculous if such were the case. Each theory arose on different occasions and for different purposes, to predict and to explain quite different phenomena. Of course one might be tempted, on this account, to treat the mental *instrumentally* , as convenient fictions. And this is precisely what Dennett does [Dennett,1971; and with regard to the net attitudes,1978, p.28]. But this temptation should be resisted. The sharing of a physical basis has nothing to do with the mental *per se* , since, as we have already had opportunity to see, the very same thing holds between macro and micro-level theories generally.

Further Confirmation for Physical Shareability

So this brings us to another reason for accepting the Physical Shareability of the Mental, and that is because the doctrine is but an instance of a more general principle which we already have reason to accept, namely, the *Microphysical Shareability of Macro Objects and Events:*

More than one object, event, or state, can have the same microphysical basis at the same time .

This more general principle is established by the previous cases of overlapping objects. The statue and its gold share a microphysical base at a particular time, as does the rope and the hammock, or the two intersecting waves. As a consequence so also will the separate events and states of those objects. But there is no temptation to treat statues, ropes, and hammocks *instrumentally* , or the events they engage in, and so impartiality demands that, on at least this account, we treat mental events in the same way.

Hence, on any intuitively plausible notion of an event, it seems possible that they can occur within the same spatio-temporal region or within the same temporal slice of an object. And if this is true, as it seems to be, then the coarse-grained view of events represented by (LQD2) must be rejected.

Now, in response, perhaps one could claim that all of the arguments given thus far only show that the identities required by (LQD2) conflict with certain inessential features of our intuitive notion of an event.

More specifically, one might concede that identifying the ball's color with its shape, or identifying my wanting to get my queen out early with my preference for a wide open game, is indeed counterintuitive but nonetheless harmless for philosophical and scientific purposes. Speaking of the ball's rotation and heating, for example, Quine says:

> I am not put off by the oddity of such identifications. Given that the ball's heating up warms its surroundings, I concede that its rotating, in this instance, warms the surroundings. I am content likewise to conclude that Sebastian's gum-chewing got him across Bologna, if it coincided with his walk. These results seem harmless to science, for they imply no causal connection between warming and rotating in general, nor between locomotion and chewing gum [Quine,1985, p.167]. [6]

I doubt, however, that we can tolerate such a revision of our common conceptual scheme. Of course, Quine is right to say that there will be no causal laws linking the aforementioned event types or properties. But trouble may arise for the respective event *tokens* , and this is our only concern here. In fact, it seems that such event identities will cause trouble for the token events, and in precisely the area Quine mentions, that of causation.

(LQD2) and Causation

Suppressing reference to times, it is true to say:

(1) The ball's bright orange color caused my eyes to fatigue.

On the other hand, it seems false to say:

(2) The ball's round shape caused my eyes to fatigue.

This appears to be good reason to reject (LQD2) since that principle requires the identification of the ball's color state with the ball's shape at that time. Yet the two differ in their causal properties, and so, by Leibniz Law, they cannot be identical. Or consider the case of my swimming. It is true to say :

(3) My swimming the Channel caused my body to become cold.

But it is clearly false to say:

(4) My shivering caused my body to become cold.

Part of the explanation, no doubt, is that the appropriate *counterfactuals* are not sustained. For I take it to be a minimal constraint on causation that true singular causal statements be counterfactually supported. On this score (1) passes the test, as does (3). But it is palpably false that:

(2*) Had the ball not been round, my eyes would not have fatigued,

and false that:

(3*) Had I not shivered, my body would not have become cold.

Of course, a defender of (LQD2) does have *something* to say about such cases. He or she could, for instance, deny that (2) and (4) are actually false. Rather, our intuitions that they are false are to be explained by appeal to pragmatic considerations. [7] On this view a statement like "the ball's round shape caused my eyes to fatigue" is true but misleading because it conversationally implies that the event in question caused the retina to fatigue in virtue of its shape, or more specifically, in virtue of a law relating the shape of the ball to certain reactions of color sensitive cells. In truth, however, that same event (which is token identical to the ball's shape) caused the fatigue of the eyes in virtue of a law relating its color to the reaction of those cells. So what is violated in such a case is the pragmatic convention that we use only the causally relevant description of an event in the context of explanation.

Now I will admit that one could explain away the apparent falseness of propositions like (2) and (4) by appeal to pragmatic considerations. Nevertheless, I am suspicious about this strategy, and I should like to indicate why I think one ought to be suspicious.

First a response made recently by Bennett [1988]. There are, it will be admitted, acceptable cases of misleading conversational implication. When I say "Jones was sober today," when he has never touched a drop, my statement is true but shamefully deceptive because it conversationally implies that he was drunk at other times, and this, due to the pragmatic conventions which govern our practice of highlighting facts in discussion. But we should hope to keep such deception at a minimum. Indeed, given the interpretation according to which statements

like (2) and (4) come out true, that my shivering really did cause my body to become cold, that the ball's being round really did cause my eyes to fatigue, it turns out that we could be misled about *an entire range of causal statements* . There is, one would think, something more than suspicious in the suggestion that deception could run so deep and apply to such a wide range of cases. As Bennett remarks: "it is one thing to say that event-causation statements can be true in what they say (about what caused what) and false in what they suggest (about why), and it is quite another to say that this obtains as often as Quinean semantics must say that it does" [1988, p.111].

The trouble with this line of response, however, is that it is not necessary that we *be* misled about these statements. [8] No one actually goes around uttering the aforementioned causal claims anyway. Hence, contrary to Bennett's remark, it is not part of the pragmatic strategy to say that it obtains very often, or even that it obtains at all. The pragmatic strategy requires only that *if* someone did utter a proposition like (2) and (4), then they would be uttering a deceptive truth.

I think, however, that a more telling response can be made. For, given the nature of events as specified by (LQD2), the pragmatic deception strategy should not even apply. That is to say, if events really are just temporal slices of objects, then, whatever we think about the status of (1) and (3), propositions (2) and (4) will turn out false because the counterfactuals are still not sustained.

Consider again the case of the ball. I think it is plausible to suppose that it could have had a slightly different shape at the time in question (it could have been deflated just prior to the time I stared at it, so that

the ball was oblong in shape, not round). But if that temporal slice of the ball had a different shape, *it would have still caused my eyes to fatigue* . Yet this means that the counterfactual (2*) is false, just as I claimed -- it is false that had the ball not been round, my eyes would not have fatigued -- and hence (2) is false as well. Thus, if I am correct, our causal statements about temporal slices are not true and misleading, but false through and through.

Perhaps a better strategy, then, is to agree that statements like (2) and (4) are false, but deny that they are *extensional* . For example, the statement:

(5) Marrying Mr. DeSalvo explains her subsequent mental collapse.

does not seem true, whereas:

(6) Marrying the Boston Strangler explains her subsequent mental collapse.

does seem true, and this, even though "Mr. DeSalvo = the Boston Strangler." The reason is that the sentential context for "... explains ..." is intensional or referentially opaque, so that the substitution of these coreferring expressions is not guaranteed to preserve the truth value of the original sentence.

Similarly, when applied to cases involving causation, this appeal to nonextensional contexts would explain the differences in truth value between (1) and (2), and also between (3) and (4), without sacrificing the

identities required by (LQD2). That is to say, the sentential context created by the term "... causes ..." is to be understood in the same way, so that the substitution of "the ball's bright orange color" for "the ball's round shape" or the substitution of "my shivering" for "my swimming" could be claimed to alter the truth value of the containing sentences, and this, even on the assumption that "the ball's bright orange color = the ball's round shape" and "my shivering = my swimming."

Indeed, some philosophers have reduced causation to a species of the explanatory relation, with the result that the opacity of explanatory contexts will necessarily carry over into statements of causation. Monroe Beardsley, for example, says that:

> To specify the cause of an event is to give a causal explanation of it, and *if* explanatory contexts are nonextensional, as many would hold, then I don't see how causal contexts could fail to be nonextesional as well [1975, p.272].

Thus, what can we say about this nonextensional treatment of our causal statements ? The basic problem, in my view, is that it appears to make causality a language or mind-dependent relation. Explanatory contexts, for example, are referentially opaque precisely because the *description* of the object or the event matters for the truth of the explanation, that is, because explanation is explanation *to a person* , and the person may not know the object or event under its description. Thus, one may not have known that Mr. DeSalvo was the Boston Strangler, which accounts for our tendency to treat (5) and (6) as differing in truth value. But causation is *not* a mind-dependent or person-relative affair. Event causation would proceed quite nicely even if there were no minds

to contemplate it and no language to speak about it.

Of course, this is not to say that mind-dependence is a necessary condition for opacity. Perhaps there are other kinds of intensional contexts which do not entail anything about minds or language. But, as Beauchamp and Rosenberg point out, it is to say that the only reason we have for thinking that causal statements are intensional is their similarity with and even their assimilation to statements of explanation [Beauchamp and Rosenberg, 1981, p.260].

Finally, on a more intuitive level, I think we should have known that (LQD2) would lead us astray in matters of causation, and for the following reason. [9] Certainly we believe that objects cause things in virtue of their properties, and that some properties of an object are relevant while other are not. An object's weight, for example, is relevant for causing an indentation when dropped, while the object's color is not.

The fine-grained view of events represented by (JK) enables us to focus in on the causally relevant properties, and thus appeal to the right events or states of an object when we cast about for the proper causes. Put in another way, (JK) permits the causal powers of things flow from just the right properties. But (LQD2), being the coarse-grained view that it is, cannot discriminate between causally relevant and causally irrelevant properties. All that (LQD2) allows us to say is that the causal powers of a thing flow from *the entire region* where the object is located, or from the *entire temporal slice* of the object, and this will include all the causally irrelevant properties which the object may exemplify at that time (causally irrelevant, of course, to the production of some particular effect -- not irrelevant to the production of others).

This, I submit, cannot lead to a correct account of causation.

Is there a middle ground between (LQD2) and (JK) ? A not-so-coarse-grained view but not-so-fine-grained view which will accommodate our talk of causation ? I think not. Monroe Beardsley and others have suggested an intermediate position according to which, for example, "my swimming" and "my swimming the Channel" are identical because the latter *entails* the former [Beardsley,1975]. Hence, even though the constitutive properties are distinct, the events remain the same in virtue of this entailment relation.

Kim would disagree, quite rightly in my view, because the causal powers of the two purported events appear to differ. But, however that may be, Jonathan Bennett has shown how "with a slight push" the intermediate position actually collapses back into a coarse-grained view [Bennett,1988, pp.120-122]. To continue with the same example, "my swimming the Channel" will entail many other things besides "my swimming." It will entail, for instance, "my being alive," and many other such things. Thus, if my swimming the Channel caused an international incident, then the intermediate position would force us to say that it was in fact my being alive, my very existence at that time, which caused the international incident !

Once again the theory appears too coarse-grained, preventing us from narrowing in on the causally relevant properties. Hence it seems that any coarse-grained view is mistaken. The basic problem is that a single spatio-temporal position or a single temporal slice of an object can sustain more than one event, each with differing causal powers. Hence a finer-grained notion is required, and the most likely candidate is the

structural view of events. This being so, and given that the structural view is incompatible with token identities, then token physicalism should be rejected.

I should like to end this chapter, however, on a more sobering note. I have attempted to show that the alternative to the structural view faces a number of difficulties. One of these difficulties, just discussed, is a problem with causation. Nevertheless, if we reject token physicalism and embrace a dualism of mental particulars, then we are faced with a problem of our own -- the problem of *mental* causation. How, exactly, can mental and physical events interact without creating a situations in which, for example, the effects are causally overdetermined ? Some hopeful suggestions have been made by Kim in his analysis of supervenient causation [see Kim,1979 and 1984a]. But the problem of mental causation is an entire project in its own right, which I cannot adequately deal with here. Hence, our conclusion about mental events and token identities should be taken provisionally, keeping in mind the time-honored problem which dualists have an obligation to answer.

Notes to Chapter IV.

[1] This is not to say that cognitive science, or any other science, actually quantifies over events in the formulation of its laws. As William Taschek pointed out to me, laws are typically written in such a way as to quantify over what we are calling the constitutive objects, and they attribute properties to those objects. Nevertheless, such laws can be *interpreted* in terms of the corresponding events which are the objects exemplifying those properties [as in Fodor,1974, p.128]. Indeed, good metaphysics may require us to quantify over events.

[2] One strategy I will avoid begins in action theory by citing cases of "level-generated" events [Goldman,1971], and then argues against (LQD2) on the basis of their spatial-overlapping. The trouble is that most level-generated events, e.g., "my pulling the trigger," "my shooting the gun," "my killing the president," etc., appear to have a distinct spatio-temporal location. The pulling of the trigger is confined to the events required to pull the trigger, the shooting of the gun involves other sets of events required to point and aim the gun, and the killing involves still yet other events, i.e., the death of the president, each of these being located in quite different places.

[3] I owe this suggestion to Gary Ebbs.

[4] Additional examples can be found in the literature. Bennett's favorite case, e.g., is that of an ox being turned over hot coals [1988, p.108]. Here the ox's rotation is one event, and its loss of rawness or being cooked is another. Yet they occur in the same region, again contrary to (LQD2).

[5] Cummins refers to this alternative as the RTI, the representational theory of intentionality [1989, pp.14 -16 and 141 ff.]. This view happens to be quite popular among philosophers who are familiar with computational psychology, and it appears to be defended by Jerry Fodor, among many others.

[6] Quine does, however, think the problem created by adverbial modification is more serious. If, e.g., the ball's rotation and heating are the same event, and the ball rotated rapidly and heated slowly, we must apparently conclude that *the same event is both rapid and slow* [Quine,1985, p.167]. I think the problem is serious, but I do not discuss it here. For a detailed treatment, see Bennett [1985]. I should also add that, in light of the adverbial problem, Quine remarks rather innocently that we could always retreat to the structural view of events to solve the problem [Quine, ibid.]. But he would be more hesitant, I think, if he traced out the ontological consequences of accepting the structural view, specifically, the consequence of a dualism of particulars, as I argued in the last chapter.

[7] I should like to thank Stephen Yablo for reminding me of these options.

[8] A point made by Taschek.

[9] Here I am endebted especially to Bennett [1988].

[10] Perhaps some philosophers who have aligned themselves with physicalism could accept this conclusion inasmuch as these irreducible mental particulars are not traditional Cartesian substances, i.e., not immaterial "souls." Indeed, reading the older mind/body literature, one gets the impression that -- brain secretions and central state identities to one side -- what these authors were really trying to avoid was a doctrine of the soul along with any religious significance which might attach to it.

CHAPTER V.

PHYSICALISM AS SUPERVENIENCE

Supervenience is a philosopher's term of art used to designate a particular determinative relation between objects, events, and properties. When applied to psychological matters it can be summed up by the simple maxim: "no mental difference without a physical difference," or more generally, "physical indiscernibility entails indisernibility complete and entire." The doctrine is intended to capture our commonsense notion that the mental is somehow determined by or dependent upon the physical. Moreover, the supervenience relation is thought to express this dependency without presupposing any controversial *identities* between the two realms (one realm of facts can determine another without the two being identical -- compare causation). Indeed, this is often thought to be its chief virtue, especially in light of the various arguments against type and token identities.

All of this is to say that psychophysical supervenience is a popular and seemingly innocuous idea. As Quine put it: "Most of us nowadays are so ready to agree to this principle that we fail to sense its magnitude" [1978, p.163]. It is one of my aims to show that the supervenience principle has considerably less magnitude than Quine and others believe it to have. But, in any case, our primary interest in supervenience arises from the fact that certain philosophers have taken it to express a

significant version of *physicalism* . More specifically, these philosophers have believed that supervenience provides not just a necessary condition, but also a sufficient condition for a physicalist view of the world [see Quine,1978; Horgan,1981; and esp. Haugeland,1983].

In the present chapter I shall present a number of reasons to show why this view is mistaken. First, there appear to be cases in which the mental does not supervene on the physical. The exact status of these non- supervenient entities, at least within psychological theory, is presently controversial. Moreover, I happen to think that a large and important range of psychological properties *are* supervenient on physical properties. But second, and what I take to be the most important point, even if all psychological entities did supervene on the physical, this by itself would not establish a significant form of ontological dependency, one which is significant enough to ground the claim that the physical is ontologically more basic than the mind. Consequently, given that the physicalist must, at a minimum, believe that physical entities *are* basic, that they *do* constitute the basic furniture of the world, then supervenience cannot be sufficient to express the physicalist view of the world.

Concepts of Supervenience

Let us begin by specifying in a more precise way the concept of supervenience which we will be discussing. Donald Davidson is credited with being the first to introduce the concept of supervenience into contemporary philosophy of mind, and he said that:

> [S]upervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect [Davidson,1970, p.214].

Here Davidson speaks of events and changes. But we should like a broader characterization of supervenience which would also include objects and their properties as well as any static features or standing conditions of those objects. Thus we shall follow Jaegwon Kim and define supervenience in this way. [1] Where A and B are classes or families of properties, A being the *supervenient* family and B its *supervenience base* , we can say that:

A supervenes on B just in case it is necessary that, for each x and each property F in A, if x has F, then there is a property G in B such that x has G, and necessarily if any y has G, it also has F [Kim,1984b, p.165].

Thus, if A supervenes on B, then every property F in A will be *determined* by some property G in B in the sense that a law of the form (x)(Gx => Fx) will obtain. Furthermore, this dependency of F on G does not preclude the possibility that other properties in B might be lawfully sufficient for F, which is just to say that there may be "alternative supervenience bases" within the family of B properties. Indeed, this is just what we should expect if mental properties are multiply realized by physical ones. Moreover, we can also explain the supervenience of events on the basis of the above definition, since the event which is x having F will supervene on the event which is x having G just in case F in the supervenient class of properties supervenes on G in the supervenience base.

In any case, given the existence of the laws required by the supervenience relation, we are able to capture the intuitive characterization of supervenience with which we began -- there cannot be a difference or change with respect to some object having property F in A without a difference or change with respect to the supervenience base properties in B. If an object which formerly had F ceases to do so, then the laws (and the principle of modus tollens) will require that the difference be reflected in a corresponding change with respect to the properties in the supervenience base.

There are other concepts of supervenience which have been discussed in the literature. Most important are those which have been defined in terms other than individuals and their properties, for example, those which use larger space-time regions or even entire worlds [see Horgan, 1982; Haugeland,1982; and Kim 1987]. But we will concentrate on the above notion, taking opportunity to mention others as the need may arise. Hence, assuming that we have an adequate understanding of the supervenience relation, let us turn to the issue about how, precisely, this relates to the doctrine of physicalism.

Physicalism as Supervenience

Donald Davidson, as I said before, introduced the concept of supervenience into contemporary philosophy of mind. But he was also the first to connect supervenience with a doctrine of physicalism. He said that his own doctrine of anomalous monism, a particular form of the token identity theory: "is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on

physical characteristics" [1970, p.214]; and later, with a more committed tone, he said that: "Although, as I am urging, psychological characteristics cannot be reduced to others, nevertheless they may be (and I think are) strongly dependent on them ... psychological concepts are *supervenient* on physical concepts" [1973, p.253]. [2]

Geoffry Hellman and Frank Thompson also add supervenience to their own particular brand of token physicalism, and, like Davidson, distinguish the result from any type reductionist program [Hellman and Thompson, 1975 and 1977]. A similar position has been more recently advanced by John Post [1987]. From our perspective, however, these authors represent a somewhat mixed view, with token identities brought in to supplement the claim about psychophysical supervenience. [3]

The first clear statement of the view I want to discuss is taken from Quine. In his paper, "Facts of the Matter," Quine said that:

> The physicalist does not insist on an exclusively corporeal ontology. He is content to declare bodies to be *fundamental* to nature in somewhat this sense: there is no difference in the world without a difference in the positions or states of bodies ... It is a way of saying that the fundamental objects are the physical objects. It accords physics its rightful place as the basic science without venturing any dubious hopes of reduction of other disciplines [1978, pp.162,163].

Thus, as long as everything supervenes on the physical, then the physical will retain its proper place as that which is ontologically fundamental, and this, even if some of the things -- not just properties but genuine objects -- happen to be noncorporeal and physically irreducible.

Terence Horgan has also expressed a similar sentiment by first observing that "the truth of token physicalism is neither necessary nor sufficient for the generality of physics" [1981, p.406], and then claiming that supervenience provides an ontological priority for the microphysical "even if certain concrete events (for instance, mental events) are not identical with physiochemical events" [1982, p.31].

The most ardent defender of this physicalism as supervenience doctrine, however, is John Haugeland. He proposes a form of psycho-physical supervenience which he describes as "a variety of physicalistic monism ... which is so weak that it entails no kind of identity theory, and yet preserves a suitable 'primacy' for the physical" [1981, p.93]. In fact, Haugeland argues at length for the view that supervenience captures all of the basic intuitions which originally motivated the identity theory [1982]. Finally, Jaegwon Kim has for some time defended a similar version of psychophysical supervenience, one cast in terms of a dualism of events [Kim,1979,1982, and 1984a]. [4] Stressing both the importance of psychophysical supervenience in the philosophy of mind and its connection to physicalist or materialist views, Kim says that:

> Acceptance or rejection of the supervenience of the mental on the physical leads to the most basic division between theories of the mind-body relation: thoeries that accept psychophysical supervenience are fundamentally materialist, and those that reject it are fundamentally anti-materialist" [Kim,1984b, p.156]

What these writers have in common, then, at least to a large extent, can be stated by the following three theses: (a) the supervenience

relation suffices to establish one set of entities as *ontologically more basic* than another; (b) this priority suffices to capture *physicalist intuitions* ; and (c) all mental entities *do* in fact supervene on the physical. I will question all three assumptions.

Exceptions to Psychophysical Supervenience

As a general doctrine about how things relate to the physical, supervenience cannot, I think, be accepted. Surely mathematical facts are an exception, at least if we are realists about mathematics. [5] But, on an intuitive level, psychophysical supervenience seems to be a much more plausible view. Psychological differences, it would seem, ought to be reflected in a difference within the physical realm; and physical similarities or indiscernibility, on the other hand, ought to guarantee the similarity or indiscernibility of the mind. Nevertheless, I think there are exceptions even here, though, again, I should add the proviso that it is presently a matter of great debate how these non-supervenient entities stand with respect to psychological theory.

Some alleged violations of supervenience I will not press. For example, those who take seriously the "inverted qualia" problem [e.g., Block and Fodor,1972], are apparently willing to countenance violations of supervenience for at least some types of mental states. For in the case of inverted qualia, individuals who are physically and functionally *equivalent* are thought to differ in the qualitative aspects of their minds, and hence the supervenience of these aspects would failto obtain. In my own view, however, it is difficult to specify the functional equivalence (including the equivalence of beliefs about one's sensations)

without making reference to these qualitative features [cp. Shoemaker,1984]. In any event, I think a more clear and straightfoward example can be presented.

Let us begin by considering a familiar twin-earth case [cp. Putnam, 1975]. Suppose we have an earthling, call him Oscar, and an exact physical replica, twin-Oscar, which happens to live on another planet. Oscar and twin-Oscar are indistinguishable in all physical respects, atom for atom, molecule for molecule, and so on. As a result, the doctrine of psychophysical supervenience would require that Oscar and twin-Oscar be psychologically indistinguishable as well. But suppose that Oscar is thinking of Vienna. Even though twin-Oscar is an exact physical duplicate, and even though Twin-Oscar might have all the same images that Oscar has when he thinks of Vienna, and so on, nevertheless, twin-Oscar does *not* think of Vienna. The reason is simply that Vienna exists on earth, not twin-earth, and twin-Oscar has had no contact of any kind with our own city of Vienna. [6]

The standard solution to this problem, of course, is to point out that the example only shows that Oscar's belief about Vienna does not supervene on the *individual's current, internal physical states* . Rather, such beliefs, often referred to as beliefs with "wide content," supervene on a wider range of physical states. As Kim observes:

> Whether a person who is in a certain *current* psychological state is thinking of Vienna is likely to depend on the person's *past* associations with that city. All this means, of course, is that thinking of Vienna does not depend on a person's current physiological state; we need to take a wider (temporally stretched) supervenience base [Kim,1979, p.49, fn.16].

Hence, if we widen the supervenience base, spatio-temporally, to include *relations* with the intentional object, in our case, Vienna, then the supervenience of Oscar's belief will presumably hold [see also Kim, 1981; and Horgan,1982]. So the trouble with twin-Oscar was simply that he and his world lacked the proper supervenience base. There was *not* the requisite physical indiscernibility, and so this is no counterexample to supervenience.

Now this is true about the twin-earth case, as described. What has been overlooked, however, is that we can make a simple adjustment within the story to provide a more interesting counterexample. For suppose that twin-earth has, not just an exact duplicate of Oscar, but an exact duplicate of *Vienna* as well -- it is not Vienna, of course, but twin-Vienna. Moreover, suppose that twin Oscar has all the same relations, all the past associations, with twin-Vienna that Oscar has with Vienna. Thus, *Oscar and twin-Oscar are physically indiscernible and related in the same way to physically indiscernible intentional objects* . Nevertheless, Oscar has a belief about *Vienna* , whereas twin-Oscar has a belief about *twin-Vienna* . The wide content differs, and thus we have an apparent violation of supervenience. [7]

The options, at least for those who defend a general view about psychophysical supervenience, are, I think, just two. First, one could propose to widen the supervenience base even further, that is, resort to what is sometimes called "global" supervenience in order to guarentee that the mental supervenes on the physical [Haugeland,1982]. Or, second, one could hold that this violation and others like them are of no real importance, since psychological theory should be concerned only with

"narrow" content, that is, the content which supervenes on an individual's internal states [see Fodor,1980; and Stitch,1978].

According to the first option, the appeal to global supervenience, the idea is to define supervenience in terms of entire worlds: any two worlds exactly alike in all physical respects must be exactly alike in all psychological respects as well. Thus, even though Oscar and twin-Oscar of W_1 do not have the same beliefs, widely construed, yet any other possible world W_2 which is physically indiscernible with respect to W_1 will have an Oscar with beliefs about Vienna and a twin-Oscar with beliefs about twin-Vienna.

The trouble with this proposal, however, is that global supervenience seems to lose the connection we intuitively think to obtain between the mental and physical realms. For example, it is consistent with global supervenience that two worlds could differ in some apparently insignificant respect, physically speaking (perhaps Oscar in W_1 differs from Oscar in W_2 merely on account of one freckle), and yet the the respective psychologies could differ enormously [see Kim,1987]. This, one would think, could not be a part of a correct account of psychological determination by the physical. Or, better yet, it cannot be a *complete* account of psychological determination by the physical. (And this is terribly important for our concerns, since, as I said before, the main point I want to establish is that the supervenience relation by itself is not sufficient to establish a significant version of physicalism.)

As for the second option, discounting the violation of supervenience because psychological theory should be concerned only with narrow content, this is at best a tendentious view. I do happen to believe that

psychological theory, specifically computational psychology, presupposes a notion of narrow content. But the question is whether psychological theory should be concerned *only* with narrow content. Burge has given case studies which point in the other direction [Burge,1986]. It suffices to say that if the physicalist is going to reject our appeal to common beliefs and other propositional attitudes, those which are individuated widely, then he or she owes us a convincing argument as to why psychological theory should be so limited.

Now Stitch and Fodor do attempt to supply the arguments, but I find them unconvincing. One of the main arguments, crudely put, is that only internal states can explain behavior, since only internal states can *cause* the behavior. Thus, apparently, we need a supervenience base for psychology which is restricted to those properties internal to the individual [cp. Kim, 1982, p.65]. But, as Burge points out, the question is not about which states do the causing (it is those located within the individual, to be sure); the question is how the internal states which do the causing are to be *individuated*. Causation is not the same thing as individuation, and Burge thinks the states in question need to be individuated widely [see Burge,1986].

Another slightly different argument is that, since Oscar and twin-Oscar will exhibit what seems to be the same behavior (for example, both might travel to their respective cities), then we need to give the same explanation in each case. As Fodor puts it: "We need, in particular, a taxonomy according to which [persons] have the *same* belief in order to explain why it is that they exhibit the same behaviors" [1980, p.239]. I take it that something like this also lies behind Stitch's

"replacement argument" [Stitch,1983, p.165]. Hence we must appeal to the narrow content, whatever it may be, which Oscar and his twin will share. The problem with this line, however, is that it seems to beg the question against those who appeal to beliefs having wide content. For, on a wide construal, it is simply not clear that Oscar and his twin *do* exhibit the same behavior -- one believes that he is traveling to Vienna, and so his intentional behavior is described as "traveling to Vienna"; whereas the other believes he is traveling to twin-Vienna, and so his behavior is described as "traveling to twin-Vienna."

Supervenience and Ontological Priority

Let us suppose, however, that all mental entities do in fact supervene on the physical. Again, I think it is true for a number of psychological states, specifically those with narrow content. But let us suppose that there are no exceptions to supervenience whatsoever. Would this be sufficient to establish the physical as *ontologically more basic* than the mental ? As far as I know, every philosopher who has written on the topic has assumed not only that the supervenience relation will establish some form of ontological dependence, but, what is most important, that it is a dependence which is significant enough to show that physical properties are ontologically more "basic" or more "fundamental" than any supervenient properties. Upon reflection, however, one would think that this assumption should be questioned.

After all, there are *other* dependency relations which do not establish any ontological priority. Effects, for example, are dependent upon their causes. Yet we do not conclude that effects are ontologically inferior

than their causes, or that causes are more basic than their effects. For instance, I have been caused by my parents. But they are not more basic on that account. Notice, also, that our position *vis-a-vis* ontological priority would not change even if I did no causing myself, and nothing causally depended on me (mirroring the asymmetrical dependence thought to obtain in the case of psychophysical supervenience). Still, I would not be inclined to say that my parents are ontologically more basic than myself.

Of course, causation is not the same relation as supervenience. Nonetheless, they are both *dependency* relations. And the point is that a dependency relation *per se* does not seem to establish one thing as being ontologically more basic than another. Hence, those who accept physicalism as supervenience owe us an explanation over and above the mere fact that mental properties depend upon the physical ones.

Now one might think that the situation would change if, as in the case of supervenience, an *entire class* of things depended on another. Imagine, for example, an entire class of things which was caused and did no causing (mirroring the dependence of entire families of properties as expressed in the definition of supervenience). Indeed, this is precisely the situation with respect to epiphenomenalism, where the mental is caused by the physical but does not cause anything else. On this view, is it not true that the one class of things is more basic, ontologically speaking, than the other ? Similarly, then, if the entire class of mental properties are dependent on the physical in the sense of being *supervenient* on the physical properties, should we not also judge that the physical is more basic, ontologically speaking ?

In response, I grant that we do happen to judge this way in the case of epiphenomenalism. But the question is *why* we do so. Is it merely because one class of things is dependent, in this case causally dependent, on the members of another class ? I think not. For we know of many classes which causally depend upon others -- artifacts, synthetic chemicals, and domesticated animals -- which seem to be no less basic, ontologically speaking, than the things which make them. To put it somewhat facetiously, am I more basic than my favorite chair because it belongs to a class of things (the artifacts) which is caused by the members of a class to which I belong ? Or, to put it differently, if this *is* the kind of priority which the physicalist has in mind, then it strikes me as being trivial and uninteresting.

But let us put the comparison with causal dependency to one side. For I think we can present a clear case in which one class or family of properties supervenes on another, and yet no ontological priority would obtain. And if this is the case, then my main point will have been established: supervenience by itself will not suffice to capture physicalist intuitions.

First let us consider what would be referred to as a case of "weak" supervenience, that is, one which will not obtain across possible worlds [see Kim,1984b]. Consider, for example, Laplace's demon. It knows all the initial conditions, all the present physical states, and all the future states of the physical world. Hence there could be no difference in the physical world without a corresponding difference which is reflected in its mind. [8] But this is just to say that the physical properties supervene on the mental properties of Laplace's demon ! Yet its mental life would

not be *more basic* from an ontological point of view. Mere knowledge, it would seem, could not establish the knower as more fundamental than what is known. In fact, the reverse seems possible, that a knower could have as its object of knowledge those things which are genuinely fundamental in the ontological sense.

Of course, Laplace's demon, we presume, does not exist in all metaphysically possible worlds. But suppose this were true. Suppose, for example, we talk about God and his omniscience. Here we would have a case of "strong" supervenience, that is, it would be true that, necessarily, any change in the physical would be reflected in God's mind. Granted, we do not have to believe that this story is actually true. It is rather a conceptual test, to find out whether the supervenience relation, by itself, will establish an interesting form of ontological priority. And my point remains. The mere supervenience of objects of knowledge upon the mental states of the knower would not suffice to establish those mental states as ontologically more basic than the objects known. Hence, if I am right, what this example shows is that there can be supervenience without ontological priority.

Now one might respond by saying that, in the above cases, there is a certain subclass of the mental states which *also depend upon the physical*. This would be true, for example, if we accepted a causal theory of knowledge, with physical states causing certain thoughts (though of course not all of the thoughts) in God or the demon's mind. Hence a subclass of the supervenience base properties in our story would actually depend, albeit *causally* depend, upon the supervenient properties. So the claim might be that the dependency which establishes

an ontological priority must be an *asymmetrical dependence in toto* , across the entire family of properties and including all forms of dependence: causal, supervenient, and whatever else is deemed relevant.

Notice, however, that if we do take this line, then we have abandoned the idea that supervenience alone suffices to establish the physical as the ontologically basic realm -- precisely the point I want to argue. For ontological priority now requires supervenience plus the absence of any other dependency relations which might hold in the opposite direction. Moreover, a different problem with this suggestion about "asymmetrical dependence *in toto* " is simply that, if it is true, then it is no longer clear that the mental *is* dependent upon the physical in the desired way; and hence it is no longer clear that the physical *is* more basic than the mental. For it could be claimed that certain physical states are also causally dependent upon the mental (for example, artifacts depend upon human design), which is just to say that the mental has causal efficacy in the world. Of course, this will not be true if we assume some form of the *identity theory* . But that just shows, again, that the supervenience relation by itself cannot establish any ontological priority for the physical.

Summary of Dissertation

What I have attempted to show is that physicalism in its various forms is a mistaken idea, and, as a result, that no doctrine of physicalism should act as a constraint upon psychological theory. Type physicalism, token physicalism, and physicalism as supervenience all face a number of difficulties which, I think, cannot be reasonably

resolved. Of course, there may be other ways to express physicalist or physicalist-related sentiments. For example, I have not addressed the eliminativist position [e.g., Churchland,1981; but cp. Horgan and Woodward,1985]. Nevertheless, the views which I have examined in the course of this study are thought by many to be the most plausible expressions of the physicalist point of view. And given their failure, we may conclude that psychological entities are indeed metaphysically autonomous *vis-a-vis* the physical sciences, which is just to say that psychology appears irreducible across metaphysical categories.

Notes to Chapter V.

[1] This is Kim's definition of "strong" supervenience. In contrast, "weak" supervenience lacks modal force (the connection between F and G is accidental), and is on that account philosophically uninteresting since accidental correlations do not capture the idea that the physical *determines* the mental. I should also mention that Haugeland is quite wrong to claim that Kim's definition presupposes token identities [Haugeland,1982, p.96]. The idea, I take it, is that *the same x* which has F (the mental property) also has G (the physical property), making x a physical entity. True enough. But we can take the variables in the definition to range over *objects* , as Kim suggested, leaving events to be constructed out of these objects (along with properties and times). Hence, even though x is a physical entity, the event which is [x,F,t] need not be identical with the event which is [x,G,t]. This will depend upon the identity of the constitutive properties, as we saw from chapter 3.

[2] Contrary to Davidson, I do not think that any interesting form of supervenience is consistent with the "anomalous" part of his monism, i.e., with the denial of all psychological and psychophysical laws [see Kim,1979]. As we have seen, Kim's relation of strong supervenience will entail one-way conditional laws from the supervenience base to the supervenient properties, although, contrary to Kim, I do not allow for the construction of infinite disjunctive properties so that two-way or biconditional laws will be forthcoming [see Kim,1984b; and cp. my chapter 2].

[3] This version of token physicalism is expressed by Hellman and Thompson's "principle of exhaustion," according to which all particulars are to be identified with the tokens of our basic physical predicates, or identified with the combinations of those tokens [Hellman and Thompson,1975, pp.553-555; and Post,1987, pp.125 and 169]. Criticism of this principle can be found in Haugeland [1983] and Levinson [1983]. It is important to realize, however, that this kind of token identity theory differs from the kind examined and rejected in chapters 3-4 of this thesis. Specifically, the token identity theory discussed there involved the identity of *events* , which is the way token identity claims were typically understood in the literature; whereas the principle of exhaustion seems to govern only the *objects* which are involved in any given event. Put differently, events are not simply the combination of objects falling under basic physical predicates, but combinations of those

objects *plus* specific properties and times (their structure is not the mereological structure of the calculus of individuals). Hence, it is perfectly feasible that one could reject token physicalism in our sense, as applied to events, and yet accept the principle of exhaustion for concrete objects.

4 It should be noted that Kim, unlike the others just mentioned, is careful not to assume that the supervenience relation will usher in a significant version of physicalism. He said: "Lest you think that an affirmative answer to the question of psychophysical supervenience automatically yields physicalism, let me remind you that G.E. Moore, to whom the thesis of moral supervenience is often attributed, was a staunch and generally effective critic of ethical naturalism ... If Moore was consistent, then by symmetry of reasoning the doctrine of psychophysical supervenience ought to be compatible with the denial of physicalism" [1982, p.52]. However, the context of this passage seems to indicate that Kim has in mind the doctrine of *type* physicalism, which other defenders of supervenience will also deny. Also, it turns out that Kim *doesn't* think Moore was consistent after all, given both the supervenience relation and his views about reduction *via* infinite disjunctive properties [Kim,1978 and 1984b].

5 Some, e.g., Quine [1978, p.162] and Post [1987, pp.166-169], simply build the exception into their definition of supervenience. Notice, however, that our intutive notion of supervenience is *trivially* satisfied in the case of mathematical objects, i.e., since there can be no change in the Platonic realm, then the conditional "there can be no difference in Plato's heaven without a difference in the physical world" will come out true. But, of course, the mathematical does not supervene on the physical, since mathematical truths do not depend upon these contingent physical facts at all.

6 Putnam' explanation would be that the meaning of 'Vienna' is determined *causally* , and twin-Oscar has no causal contact with the city. Burge [1979 and 1986] would supply a different explanation in terms of the different *social roles* which the term has in the communities.

7 One might object that the relevant physical indiscernibility does not obtain for the following reason. Granted, Oscar and his twin are related in the same way to their respective

intentional objects. But, aside from *these* relations, there are other relational properties which make a physical difference between the supervenience bases; e.g., Vienna and Twin-Vienna are physically discernible, say, because one of them is closer to the spiral nebula than the other. True enough. However, (a) on an intuitive level, these other relations to far away objects seem *irrelevant* to the mental states of Oscar and his twin, and (b) by incorporating more and more of these external relations we come very close to adopting the "global" supervenience view, which I will discuss shortly.

8 We can also suppose the requisite asymmetry, that there *can* be a difference in the demon's mind without a corresponding difference in the physical world (perhaps it reflects upon itself as well).

BIBLIOGRAPHY

BIBLIOGRAPHY

Armstrong, D.M. (1973). *Belief, Truth, and Knowledge* . London: Cambridge Press.

_____. (1978a). *Nominalism and Realism* . Cambridge Press.

_____. (1978b). *A Theory of Universals* . Cambridge Press.

Armstrong, D.M. and Malcom, Norman. (1984). *Consciousness & Causality: A Debate on the Nature of Mind* . London: Basil Blackwell.

Beardsley, Monroe. (1975). "Actions and Events: The Problem of Individuation," *American Philosophical Quarterly* , vol.12, 263-276.

Beauchamp, Tom L. and Rosenberg, Alexander. (1981). *Hume and the Problem of Causation* . Oxford: Oxford University Press.

Bennett, Jonathan. (1985). "Adverb-Dropping Inferences and the Lemmon Criterion," in Lepore and McLaughlin, *Actions and Events: Perspectives on the Philosophy of Donald Davidson* . London: Basil Blackwell.

_____. (1988). *Events and their Names* . Indianapolis: Hackett Publishing.

Block, Ned. (1978). "Troubles with Functionalism," in Ned Block, ed., *Readings in Philosophy of Psychology* , vol.1. Cambridge Mass: Harvard University Press,1980.

_____. (1980). "What is Functionalism ?," in Ned Block, ed., *Readings in Philosophy of Psychology* , vol.1.

Block, Ned. and Fodor, Jerry. (1972). "What Psychological States Are Not," *Philosophical Review* vol.81; rpt. in Ned Block, ed., *Readings in Philosophy of Psychology* ,vol.1.

Boyd, Richard. (1980). "Materialism without Reduction: What Physicalism Does Not Entail," in Ned Block, ed., *Readings in Philosophy of Psychology* .

Brand, Myles. (1976). "Particulars, Events, and Actions," in Brand and Walton, eds., *Action Theory* . Dordrecht: Reidel,1976.

Bunzl, Martin. (1979). "Causal Overdetermination," *Journal of Philosophy* , vol.76, 134-150.

Burge, Tyler. (1975). "Mass Terms, Count Nouns, and Change," *Synthese* , vol. 31, 459-478.

_____. (1979). "Individualism and the Mental," *Midwest Studies* , vol.4, 73-121.

_____. (1986). "Individualism and Psychology," *Philosophical Review* , vol.95, 3-45.

Chisholm, Roderick. (1970). "Events and Propositions," *Nous* , vol.4, 15-24.

_____. (1971). "States of Affairs Again," *Nous* , vol.5, 197-189.

Churchland, Patricia. (1987). "Epistemology in the Age of Neuroscience," *Journal of Philosophy* , vol.84, 544-553.

Churchland, Paul. (1981). "Eliminative Materialism and the Propositional Attitudes," *Journal of Philosophy* , vol.78, 67-90.

Clark, Romane. (1970). "Concerning the Logic of Predicate Modifiers," *Nous* , vol. 4, 311-335.

Cummins, Robert. (1986). "Inexplicit Information," in Brand and Harnish, ed., *The Representation of Knowledge and Belief* . Tucson: University of Arizona Press.

_____. (1989). *Mental Representation* . M.I.T. Press.

Davidson, Donald. (1980). *Essays on Actions and Events* . Oxford Press.

_____. (1967a). "The Logical Form of Action Sentences," rpt. in *Essays* .

_____. (1967b). "Causal Relations," rpt. in *Essays* .

_____. (1969). "The Individuation of Events," rpt. in *Essays* .

_____. (1970). "Mental Events," rpt. in *Essays* .

_____. (1973). "The Material Mind," rpt. in *Essays* .

_____. (1974). "Psychology as Philosophy," rpt. in Essays .

_____. (1985a). "Reply to Quine on Events," in Lepore and McLaughlin, eds., *Actions and Events: Perspectives on the Philosophy of Donald Davidson* . Basil Blackwell,1985.

_____. (1985b). "Replies," in Bruce Vermazen and Merrill Hintikka, eds., *Essays on Davidson: Actions and Events* . Oxford, Clarendon Press,1985.

Dennett, Daniel. (1978). *Brainstorms* . M.I.T. Press.

_____. (1971). "Intentional Systems," rpt. in *Brainstorms* .

_____. (1976). "A Cure for the Common Code," rpt. in *Brainstorms* .

_____. (1978). "Reply to Arbib and Gunderson," in *Brainstorms* .

Enc, Berent. (1983). "In Defense of the Identity Theory," *Journal of Philosophy* , vol.80, 279-298.

Feigl, Herbert. (1958). "The 'Mental' and the 'Physical,'" in H. Feigl and G. Maxwell, eds.,

Minnesota Studies in the Philosophy of Science, vol.2, 370-497.

Feldman, Fred. (1980). "Identity, Necessity, and Events," in Ned Block, *Readings in Philosophy of Psychology*, vol. 1. Harvard University Press.

Feyerabend, Paul. (1963a). "Materialism and the Mind/Body Problem," *Review of Metaphysics*, vol.17, 49-67.

_____. (1963b). "How to Be a Good Empiricist -- A Plea for Tolerance in Matters Epistemological," rpt. in Baruch Brody, ed., *Readings in the Philosophy of Science*. Prentice-Hall, 1970.

Fodor, Jerry. (1974). "The Special Sciences (Or: The Disunity of Science as a Working Hypothesis)," in *Representations*.

_____. (1980). "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology," in *Representations*. M.I.T. Press.

_____. (1981). "Something of the State of the Art," in *Representations*.

_____. (1986). "Why Paramecia Don't Have Mental Representations," *Midwest Studies*, vol.10, 3-23.

_____. (1987). *Psychosemantics*. M.I.T. Press.

Gardner, Howard. (1985). *The Mind's New Science*. Basic Books, Inc.

Goldman, Alvin. (1971). "The Individuation of Events," *Journal of Philosophy*, vol.68, 761-774.

Hacking, Ian. (1983). *Representing and Intervening*. Cambridge University Press.

Haugeland, John. (1982). "Weak Supervenience," *American Philosophical Quarterly*, vol. 19, 93-103.

_____. (1983). "Ontological Supervenience," *Southern Journal of Philosophy*, vol. 22, Supplement, 1-12.

Hellman, Geoffry. and Thompson, Frank. (1975). "Physicalism: Ontology, Determination, and Reduction," *Journal of Philosophy*, vol.72, 551-564.

_____. (1977). "Physicalist Materialism," *Nous*, vol.11, 309-345.

Hempel, Carl. (1935). "The Logical Analysis of Psychology," rpt. in Ned Block, ed., *Readings in Philososphy of Psychology*, vol.1.

_____. (1966). *Philosophy of Natural Science*. Prentice-Hall.

Horgan, Terence. (1978). "The Case Against Events," *Philosophical Review*, vol. 87, 28-47.

_____. (1980). "Humean Causation and Kim's Theory of Events," *Canadian Journal of Philosophy* vol. 10, 663-679.

_____. (1981). "Token Physicalism, Supervenience, and the Generality of Physics," *Synthese*, vol. 49, 395-413.

_____. (1982). "Supervenience and Microphysics," *Pacific Philosophical Quarterly*, vol.63, 29-43.

Horgan, Terence. and Tye, Micheal. (1985). "Against the Token Identity Theory," in Lepore and McLaughlin, *Actions and Events: Perspectives*.

Horgan, Terence. and Woodward, James. (1985). "Folk Psychology is Here To Stay," *Philosophical Review*, vol. 94, 197-226.

Hull, David. (1974). *Philosophy of Biological Science*. Prentice-Hall.

Kemeny, J.G. and Oppenheim, P. (1956). "On Reduction," *Philosophical Studies*, vol.7, 6-17.

Kim, Jaegwon. (1966). "On the Psycho-Physical Identity Theory," *American Philosophical Quarterly*, vol.3, 227-235.

_____. (1969). "Events and their Descriptions: Some Considerations," in Nicholas Rescher, ed., *Essays in Honor of Carl G. Hempel*, Dordrecht Reidel, 1969.

_____. (1972). "Phenomenal Properties, Psychophysical Laws, and the Identity Theory," *Monist*, vol.56, 177-192.

_____. (1973). "Causation, Subsumption, and the Concept of Event," *Journal of Philosophy*, vol. 70, 217-236.

_____. (1976). "Events as Property Exemplifications," in Brand and Walton, *Action Theory*.

_____. (1978). "Supervenience and Nomological Incommensurables," *American Philosophical Quarterly*, vol.15, 149-156.

_____. (1979). "Causality, Identity, and Supervenience in the Mind-Body Problem," *Midwest Studies*, vol.4, 31-49.

_____. (1982). "Psychophysical Supervenience," *Philosophical Studies*, vol.41, 51-70.

_____. (1984a). "Epiphenomenal and Supervenient Causation," *Midwest Studies*, vol.9, 257-270.

_____. (1984b). "Concepts of Supervenience," *Philosophy and Phenomenological Research* vol.45, 153-176.

_____. (1985). "Psychophysical Laws," in Lepore and McLaughlin, eds., *Actions and Events: Perspectives* .

_____. (1987). "'Strong' and 'Global' Supervenience Revisited," *Philosophy and Phenomenological Research* , vol.48,

Lemmon, E.J. (1967). "Comments on D. Davidson's 'The Logical Form of Action Sentences,'" in Nicholas Rescher, ed., *The Logic of Decision and Action* . University of Pittsburgh Press, 1967.

Levinson, Jerrold. (1983). "A Thousand Entities: Comments on Haugeland's 'Ontological Supervenience,'" *Southern Journal of Philosophy* , vol.22, Supplement, 13-17.

Lewis, David. (1973). "Causation," in E. Sosa, ed., *Causation and Conditionals* . Oxford Press.

_____. (1978). "Mad Pain and Martian Pain," in Ned Block, ed., *Readings in Philosophy of Psychology* , vol.1.

Lombard, Lawrence. (1979). "Events," *Canadian Journal of Philosophy* , vol.9, 425-460.

_____. (1986). *Events: a Metaphysical Study* , Routledge & Kegan Paul.

McLaughlin, Brian. (1985). "Anomalous Monism and the Irreducibility of the Mental," in Lepore and McLaughlin, eds., *Actions and Events: Perspectives* .

Nagel, Ernest. (1961). *The Structure of Science* , Harcourt, Brace & World Inc.

Nagel, Thomas. (1965). "Physicalism," *The Philosophical Review* , vol.74, 339-356.

Nelson, John. (1985). "Physical Properties," *Pacific Philosophical Quarterly* , vol.66, 268-282.

Penfield, W. and Roberts, L. (1959). *Speech and Brain Mechanisms* . Princeton University Press.

Plantinga, Alvin. (1962). "Comments," in Hector-Neri Castaneda, ed., *Intentionality, Minds, and Perception* , Wayne State U. Press,1966.

Popper, Karl. and Eccles, John. (1977). *The Self and Its Brain* . Routledge & Kegan Paul.

Post, John F. (1987). *The Faces of Existence: An Essay in Nonreductive Metaphysics* . Cornell University Press.

Putnam, Hilary. (1967). "The Nature of Mental States," in *Mind, Language and Reality* . London: Cambridge University Press, 1975.

_____. (1973). Philosophy and our Mental Life," in *Mind, Language and Reality* .

_____. (1975). "The Meaning of 'Meaning,'" in *Mind, Language and Reality*.

_____. (1988). *Representation and Reality*. Cambridge Mass: M.I.T. Press.

Pylshynyn, Zenon. (1984). *Computation and Cognition*. M.I.T. Press.

Quine, W.V.O. (1960). *Word and Object*. M.I.T. Press.

_____. (1969). "Natural Kinds," in *Ontological Relativity & Other Essays*. N.Y: Columbia University Press.

_____. (1978). "Facts of the Matter," *Southwestern Journal of Philosophy*," vol. 9, 155-169.

_____. (1985). "Events and Reification," in Lepore and McLaughlin, *Actions and Events: Perspectives*.

Richardson, Robert. (1979). "Functionalism and Reductionism," *Philosophy of Science*,

Rosenberg, Alexander. (1978). "The Supervenience of Biological Concepts," *Philosophy of Science*, vol.45, 368-386.

_____. (1985). "Davidson's Unintended Attack on Psychology," in Lepore and McLaughlin, eds., *Actions and Events: Perspectives*.

Sanford, David. (1970). "Locke, Leibniz, and Wiggins on Being in the Same Place at the Same Time," *Philosophical Review*, vol.79, 75-82.

Schaffner, Kenneth. (1967). Approaches to Reduction," *Philosophy of Science*, vol.34, 137-147.

Shoemaker, Sydney. (1970). "Wiggins on Identity," *Philosophical Review*, vol.79, 529-544.

_____. (1980). "Causality and Properties," in P. Van Inwagen, ed.,*Time and Cause*. D. Reidel.

_____. (1981). "Some Varieties of Functionalism," rpt. in J.I. Biro and Robert Shahan, eds., *Mind, Brain, and Function*. University of Oklahoma Press,1982.

_____. (1984). "Absent Qualia are Impossible -- A Reply to Block," in *Identity, Cause, and Mind*. Cambridge University Press.

Sklar, Lawrence. (1974). "Thermodynamics, Statistical Mechanics and the Complexity of Reductions," in R.S. Cohen et al., eds., *Philosophy of Science Association* (*PSA*), D. Reidel, 1976.

Smart, J.J.C. (1959). "Sensations and Brain Processes," *Philosophical Review*, vol.68.

Sperry, R.W. (1974). "Lateral Specialization in the Surgically Seperated Hemispheres," in Schmitt and Worden, eds., *The Neurosciences Third Study Program* . M.I.T. Press.

Stitch, Stephen. (1978). "Autonomous Psychology and the Belief-Desire Thesis," *Monist* , vol.61, 573-591.

_____. (1983). *From Folk Psychology to Cognitive Science* . M.I.T. Press.

Suppes, Patrick. (1985). "Davidson's Views on Psychology as a Science," in Bruce Vermazen and Merrill Hintikka, eds., *Essays on Davidson: Actions & Events* . Oxford, Clarendon Press.

Teller, Paul. (1983). "Comments on Kim's Paper," *Southern Journal of Philosophy* , vol.22, Supplement, 57-61.

Thalberg, Irving. (1980). "Can We Get Rid of Events ?," *Analysis* , vol.40, 25-31.

Wiggins, Richard. (1967). *Identity and Spatio-Temporal Continuity* . Basil Blackwell.

_____. (1968). "On Being at the Same Place at the Same Time," *Philosophical Review* , vol. 77, 90-95.

Wilson, Mark. (1985). "What is this Thing Called 'Pain' ? -- the Philosophy of Science Behind the Contemporary Debate," *Pacific Philosophical Quarterly* , vol.66, 227-267.

Wilson, N.L. (1974). "Facts, Events, and Their Identity Conditions," *Philosophical Studies* , vol.25, 303-321.

Winograd, T. (1981). "What Does it Mean to Understand a Language?," in D. Norman, ed., *Perspectives on Cognitive Science* . Norwood, N.J., Ablex, 231-263.

Zangwill, O.L. (1960). *Cerebral Dominance and its Relation to Psychological Function* . Edinburgh: Oliver & Boyd.