Current Issues in the Mind/Body Problem

by

Israel Krakowski

B.S., Brooklyn College
(1972)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS OF THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1981

Signature of Author _____
Department of Linguistics and Philosophy
March 9, 1981

Certified by _____
Paul Horwich
Thesis Supervisor

Accepted by _____
Judith Thomson
Chairperson, Departmental Graduate Committee

CURRENT ISSUES IN THE

MIND/BODY PROBLEM

by

Israel Krakowski

Submitted to the Department of Linguistics and Philosophy
on March 9, 1981 in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy
in Philosophy

ABSTRACT

I discuss what I take to be the strongest recent arguments for and against mind/body identity. On the pro side I discuss Lewis' views, and on the con Kripke's. The discussion of Kripke takes the form of an examination of the major objections which have been made against his views.

The conclusion of my discussion of Lewis is that he has not adequately defended his claims of type/type identity. The, more tentative, conclusion of my discussion of Kripke is that his arguments by and large escape unscathed from the challenges offered. Since Kripke's arguments are directed at token/token views also, they will apply to Lewis' even if it is construed as such.

Thesis Supervisor: Dr. Paul Horwich
Title: Assistant Professor of Philosophy

ACKNOWLEDGMENT

I would like to express my extreme gratitude and indebtedness to all my family and friends who supported me throughout the writing of this thesis. My greatest debt is to my wife, Ricky, without whose support and love, this thesis could not have been completed. I also owe much to my son, Moshe, daughter, Tichye-Leah, and my parents.

I wish also to express my intellectual indebtedness to Ed McCann, Paul Horwich, and all the members of M.I.T.'s philosophy department whose discussion and thoughts have found their way, one way or another, into this thesis. Ned Block, in particular, has been, over these years, of great help, and intimately involved in the development of the thesis.

Finally, I would like to dedicate this thesis to my daughter, Tichye-Leah, who came into existence almost simultaneously with the completion of this thesis; though it must be admitted her gestation period was somewhat shorter.

BIOGRAPHY

I received my B.S. in mathematics and philosophy from Brooklyn College in January of 1972. Currently, I am working as an actuary for CNA Insurance in Chicago. I am married and have two children.

TABLE OF CONTENTS

Mind, n. A mysterious form of matter secreted by the brain. Its chief activity consists in the endeavor to ascertain its own nature, the futility of the attempt being due to the fact that it has nothing but itself to know itself with.

Pain, n. An uncomfortable frame of mind that may have a physical basis in something that is being done to the body, or may be purely mental, caused by the good fortune of another.

Ambrose Bierce, The Devil's Dictionary.

INTRODUCTION

The mind/body problem, one of philosophy's most ancient and venerable, continues to be--within the Anglo-American tradition--a stubborn source of philosophical controversy and perplexity. It is by and large true that most contemporary philosophers are materialists of one sort or another. There remains, nevertheless, a vociferous minority which, though it shares the same intellectual roots, does not hew the materialist party line.

Actually, when one examines the views of contemporary philosophers strictly on the basis of their content, those views are so varied, and the questions which they are responses to so diverse, that it often seems that there exists a continuous spectrum of views. In addition, some views, e.g., functionalism on many of its construals, are arguably "orthogonal" to the traditional disputes. Muddying the waters yet further is the unfortunate fact that prcpounders of many of these positions are often themselves unclear about which questions they wish to answer.

The present thesis is concerned with the mind/body question viewed from a metaphysical perspective. I am primarily concerned with such questions as, "Are persons identical to bodies?", and, "Are all mental states (events), particulars and universals, identical to physical states

(events)?". My main goal is to evaluate what I take to be the most forceful current arguments for and against mind/body identity: Lewis on the pro side, and Kripke on the con.

I choose Lewis on the pro side because I consider his position, as do many others, the most defensible and clearly stated one. It is a position which seems to easily handle many of the more recalcitrant problems besetting most other contemporary formulations of type/type materialism.

The first half of this thesis presents Lewis' views and considers various objections. Most of these objections are seen to have possible responses; but I show that making those responses forces Lewis into a position which diverges from the views and motivations of the standard type/type materialist. My attack on Lewis' position does not, of course, show that every variant of type/type physicalism (in this thesis I shall use the terms 'physicalism' and 'materialism' interchangeably) is vulnerable to similar critiques. To the extent, however, that we view his formulation as the most inciteful attempt to date to deal with the problems associated with (type/type) physicalism, our negative conclusion implicates, albeit indirectly, (type/type) physicalism generally. However, nothing I say in this section would affect Lewis' views if construed as a

token/token doctrine.

The second half of this thesis discusses the recent antimaterialist arguments of Kripke--who has resurrected and put into modern garb Descartes' arguments. Kripke's arguments are by now so well known and have generated so much discussion that there would be no point in recapitulating them here in detail. Instead, I shall assume that the reader is familiar with the details; and start right in, in the second half of the thesis, by considering, in individual sections, three of the strongest objections to his views. I show that none of them are compelling (though some are more persuasive than others). Though such a sampling can hardly claim to be exhaustive, these views were chosen for their comparative strength. Therefore, my tentative conclusion must be that Kripke's arguments presently stand.

The overall result of my discussion is that there presently exists a cogent unanswered objection to materialism. And since this objection is directed to both token/token and type/type physicalism, it will apply to Lewis, even if he is construed as only a token/token physicalist. Thus, I conclude that given the present state of philosophical discussion, as exemplified by the views I consider, the anti-materialist has the stronger position, especially with reference to type/type claims. It may not be checkmate, but it is check.

LEWIS

The first part of this thesis will discuss David Lewis' views on the nature of psychological (mental) and physical states and properties. Lewis is a physicalist in the tradition of those--most notably, and ambitiously, David Armstrong[1]--who argue that physicalism follows from the functional definability of mental terms and some true empirical assumptions. The interest in Lewis' views lies in their constituting the most precise formulation to date of this sort of argument for physicalism, indeed of any sort to my knowledge.[2]

Lewis is not concerned, as Armstrong for the most part is, with specific analyses of particular mental terms; he is concerned to state precisely and consistently what that view is which Armstrong's analyses purport to be particular instances of.

My discussion will be divided into four sections. The first presents what I take to be the most prominent problem with Lewis' views; and Lewis' response--which is

---

[1]Armstrong's views are most fully expounded in A Materialist Theory of Mind. Both Lewis and Armstrong suggest that their translational programs are the same. Rosenthal ("Mentality and Neutrality," JP Vol. LXXII, No. 13, July 13, 1976) suggests that Armstrong's and Lewis' views are not quite as similar as they make out. He also points out that to the extent that they are not, Lewis' views are clearly preferable.

[2]Thus, Smart calls Lewis' views "by far the best, most sophisticated and convincing" ("Further Thoughts on Identity Theory," Monist LVI, No. 2, April 1972, Pg. 162).

seen to be inadequate. The second section offers two more objections to Lewis, one of them Kripke's. Lewis' replies are given, and his views are presented in more detail. Section three examines some of the ramifications of the views presented in two. Section four is more general, concerning itself with an analysis of Lewis' construal of theoretical definition and reduction. What should be a paradigm for a term defined according to Lewis: 'Gene', is considered; other facets of Lewis' more general views are also discussed.

The conclusion we shall draw is that while it is true that Lewis' views constitute a precise formulation, formally elegant and internally coherent, the virtues are bought by having those views, when fully elucidated, actually be quite different from standard type/type physicalism, both in detail and motivation.

SECTION I

I

The thesis Lewis defends is "the hypothesis that--not necessarily but as a matter of fact--every experience is identical with some physical state."[3] One should keep in mind that Lewis' thesis purports to identify the relevant universals,[4] not(only) particulars. His argument informally put is:

---

[3]David K. Lewis, "An Argument for the Identity Theory," in Materialism and the Mind-Body Problem, David M. Rosenthal, ed.

The above quote is somewhat misleading; as it stands, it seems susceptible to attack by well-known arguments of Kripke's. When we come to a more precise formulation of Lewis' views, we shall see that it does not, in fact, succumb to these arguments. Indeed, it is one of the virtues of Lewis' views--especially when compared with other "contingent identity theorists"--that one can see immediately how he would meet Kripke's challenge.

[4]About the nature of those things Lewis means to identify, he has the following to say:

> Experiences here are to be taken in general as universals, not as abstract particulars. I am concerned, for instances, with pain, an experience that befalls many people at many times; or with pain of some definite sort, and experience which at least might be common to different people at different times. Both are universals, capable of repeated instantiation. The latter is a narrower universal than the former, as crimson of some definite shade is narrower than red, but still a universal. I am not concerned with the particular pain of a given person at a time . . . . (Lewis, Op. Cit., Pg. 162, Fn. 1)

This is not to say much about the nature of universals, but for the purposes of this section, it is enough. It is Lewis' contention that these universals (i.e., mental ones like pain) are physical, which gets him into trouble. If he would have restricted himself to what is called 'token/token identity'--the claim that each particular mental event is identical with some particular physical event--the problems I suggest for Lewis' view would not have arisen.

(1) Mental state M = the occupant of causal role R (by the definition of M)
(2) Neural state N = the occupant of causal role R (by the physiological theory)
(3) Mental state M =$_5$ neural state N (by the transivity of =)

Premise (1) is supposedly definitional; it follows from the definition of mental terms--in this case, M.[6] Psychological terms referring to mental state (events, properties, etc.) are, as are theoretical terms, definable in terms of their most typical causes and effects.

(Fn. 4 continued)

Lewis also writes, "States also are to be taken in general as universals. I shall not distinguish between processes, events, phenomena, and states, in a strict sense" (Ibid.). These are all on Lewis' view (I believe), a species of property. I shall go into Lewis' views on the nature of universals in more detail in section two. For now, and unless otherwise specified, I shall follow Lewis and not distinguish among any of these when I speak of universals, comments about which should be taken to apply equally well to properties, states, experiences, etc. (that is, pain as well as the property of being in pain as well as the experience of pain). I shall be using the relevant terms ('property', 'state', 'experience') interchangeably (where I wish to distinguish them, I shall do so explicitly).

We should also note that by 'physical', Lewis seems to mean something like, 'referred to by some term of physics', or perhaps, 'referred to by some term which is reducible to a physical term'.

[5]This is Lewis' own summary from "Psychological and Theoretical Identification," in Australian Journal of Philosophy, Vol. 50, No. 3, Dec. 1972.

[6]Lewis often construes letters--e.g., 'M', in 'by the definition of M'--as referring to themselves. When discussing Lewis, or where it is clear from context, I shall do so also; otherwise, I will follow standard conventions, using corner quotes, single quotes to mention letters and words, and double quotes for longer quotes and scare quotes. Green letters should be construed as metalinguistic variables.

One forms such a definition by first collecting all the platitudes, generalizations, and causal statements of common sense psychology. One then conjoins all these to form a 'theoretical postulate' which we write: T(t).[7] 'T' here represents the causal role which the psychological states, events, etc., must satisfy; and 't' is shorthand for 't_1 . . . t_n' which are construed as names of psychological states and events. From this postulate, one forms a modified Ramsey sentence, by first replacing all the psychological terms (i.e., 't') with variables and placing corresponding existential quantifiers out front, getting: $\exists x Tx$, and then forming: $\exists! x Tx$, the modified Ramsey sentence (which is logically equivalent to $\exists y \forall x(Tx \leftrightarrow x)$). One then forms the desired identity (which is taken to be definitional: $\iota x Tx = t$,[8] (which is equivalent to $(\exists! x Tx \rightarrow Tt) \wedge \neg \exists! x Tx \rightarrow t = *)$

---

[7] I shall henceforward take it that there is only one psychological term 't' to be defined; and that there is only one variable 'x' which replaces it. This simplification will not affect the main issues. It should be pointed out that one of the virtues of Lewis' analysis (as opposed to say Armstrong's) is that it allows for mental terms to be defined in a way that mentions other mental entities, as Rosenthal points out (Op. Cit.).

[8] Lewis gives two everyday examples of such definitions. The first ("An Argument for the Identity Theory," Pg. 109) is that the definitive character of a (cylindrical) lock's being unlocked is the "syndrome of its most typical causes and effects, namely that setting the combination typically causes the lock to be unlocked and that being unlocked typically causes the lock to unlock when gently pulled . . . alignment of the bolts occupies precisely the causal role that we ascribed to being unlocked by analytic necessity, as the definition characteristic of being unlocked (for these locks)."

It is important to realize that the sort of definition just summarized is intended to apply to two cases. First, to theoretical terms generally (as adumbrated in "How to Define Theoretical Terms"), T(t) in such cases formulates a bona fide theory in which 't' denotes the new theoretical terms introduced by the theory.

Secondly, it applies to our everyday, common sense

Fn. 8 continued)

The other example is that of a detective who introduces characters called only 'X', 'Y', 'Z', as the agents in a complicated tale of intrigue and murder (without further saying who these characters are). These terms are then analytically defined as referring to those three people fulfilling the roles presented in the detective's story. If there is such a unique triple, then X, Y, Z, are those three.

For an example from folk psychology, I shall simplify tremendously. Say that the common beliefs about pains are that they are often caused by the cutting of skin and that they often cause shouting behavior. (To be really a proper example, I should mention other psychological states in my beliefs about pain; 'pain' itself would not be defined in isolation, but rather as one term among many in an enormous postulate which constitutes folk psychology. As long as we keep this in mind, things will be clearer if we simplify.) Our postulate 'T(t)' will now be: 'pain is often caused by cutting skin & pain often causes shouting behavior', where 'pain' is our 't'. The Ramsey sentence of this postulate, '$\exists x$ (x is often caused by cutting skin & x often causes shouting behavior)', asserts that something satisfies the causal role T(), and the modified Ramsey sentence '$\exists!x$ (x is often caused by cutting skin & x often causes shouting behavior)', asserts that there is just one thing which does so. Our definition of 'pain', is 'ιx(x is usually caused by cutting skin& x usually causes shouting behavior)'. Thus, 'pain' refers, by definition, to the unique entity, if there is such, which satisfies the predicate '__ often caused by cutting skin & __causes shouting behavior'.

For a more comprehensive treatment, see, "How to Define Theoretical Terms," JP Vol. LXVII, No. 13, June 1970. The technical details of Lewis' treatment are somewhat more complex, but not crucial here.

mental terms.[9] Thus, if we think of our everyday psychological beliefs as a "folk" theory and our psychological terms as having been introduced as theoretical terms, we shall get the appropriate definitions. To be sure, psychological terms were, in fact, never introduced this way; the claim that they were is, says Lewis, a myth--but a convenient myth, for it entails that our everyday psychological terms have the meanings they in fact do.

Taken by itself, Lewis' analysis of the meaning of our everyday mental terms is on shaky grounds; indeed it is, I believe, probably incorrect. Nor does Lewis actually argue in its defense. What he says is that it captures behaviorism's insights: viz., that there is an analytic component in the relation between states and typical causes and effects, without behaviorism's defects, most of which are circumvented by actually identifying mental states with physical (neurophysiological) states. But these considerations are not intended as a complete reasoned defense powerful enough to convince the sceptic; they are meant, rather, to capture the motivation behind Lewis' definitional program for those willing to be convinced (though in fairness to Lewis, we must say it is hard to think of what else one might adduce in support of one's

---

[9] To allow that some of our everyday beliefs might be false and mental terms refer, Lewis suggests that we take instead of a simple conjunction of all our beliefs, "a disjunction of all conjunctions of most of them." This modification is not important here.

meaning analyses of everyday terms).

Secondly, even if Lewis' meaning analyses are correct, there is still the objection, made by Nagel against Armstrong, that, "even if some form of materialism is true, it will not automatically be expressible in the framework of common sense psychology . . . . The psychology of common sense, embodied in the ordinary concepts of behavior, desire, sensation, perception, and emotion, and so forth, is not a scientific theory. The mental states for which Armstrong offers causal analyses are picked out by a system which has evolved naturally, and whose form may depend significantly on its extra-scientific functions."[10] Or, in other words, it is possible that pain may go the way of phlogiston; indeed, it is not only possible, but also probable by the lights of recent philosophers.[11]

When, however, we take into account the subsumtion of everyday psychological definition to theoretical definition in general, these problems lose their force. Consider first the second problem. Lewis is certainly willing to grant that pain may go to the way of phlogiston (an example he himself uses). In both folk psychology and bona fide theories, if the theoretical postulate is not realized or if it is multiply realized, then, claims Lewis, the theory is false

---

[10] "Armstrong on the Mind," PR LXXIX, 1970.

[11] See "Why Robots Can't Feel Pain," by D. Dennet, in Brainstorms (U.S.: Bradford Books, 1978), for an argument to this effect in the case of pain.

and the terms therein (e.g., 'pain') denotationless. But the fact that pain is not identical to any physical state because it does not exist hardly counts as a threat to Lewis' physicalism. If some other scientific psychological theory evolved out of our folk theory, then its theoretical terms would get defined as are all theoretical terms, and the argument for physicalism could proceed along normal lines.

A similar response can be made to the first point. Even if Lewis' definitions of everyday psychological terms are incorrect, he still has his general account of theoretical definition which he may apply to whatever psychological theory(ies) ultimately develops; and his physicalism can reduce these to physical terms.

So Lewis' physicalism gains considerable credence by its subsumtion of the case of psychological definition under the more general class of theoretical definitions. And if one is dubious of Lewis' analysis of everyday psychological terms, he can let, without any real weakening of Lewis' argument, talk of, e.g., 'pain', be construed as standing proxy for talk of terms of some future psychological theory.

Before continuing, let me turn cursorily to the second premise of Lewis' argument.

That, claims Lewis, follows from the standard belief in the dogma that all sciences will, ultimately, be reducible to physics (`a la Putnam and Oppenheim); that all phenomena are ultimately explainable by physics. Thus: "My second

premise does not rule out the existence of nonphysical phenomena . . . . It only denies that we need ever explain physical phenomena by nonphysical ones."[12] We need never advert to nonphysical entities in our explanations of the nature of the world.

I do not wish to entangle myself here in the intricate problems surrounding the notion of explanation. We might point out, however, that we could grant that if physics could explain all the phenomena there is to explain, then there might be good reason for identifying the objects, states, etc., of psychology (and other special sciences) with physical objects, states, etc. But there do seem to be good reasons, made most forcefully in a series of recent articles by Putnam, for thinking that no such comprehensive explanatory power is to be expected of physics. One such reason (in the case of psychology) is that different physical properties are, it is likely, correlated with any given psychological property (or state).[13]

Questions of explanation aside, the fact that most likely there are many physical states correlated with any

---

[12]"An Argument for the Identity Theory," Pg. 169.

[13]In fact, Putnam contends that even if it turned out that there was only one physical property corresponding to a (or every) psychological property, it would still not follow that an explanation of the relevant phenomena could be given on the basis of this physical property. For further expatiation and examples, see, "Reduction and the Nature of Psychology," Cognition 2(1), pp. 131-146, and "Philosophy and Our Mental Life," in Mind, Language, and Reality: Philosophical Papers Vol. 2, pp. 291-303; especially Pg. 293.

one given psychological state (in different species or in different people) poses a serious problem for Lewis; for it would appear that we are not justified in identifying psychological states and properties with any one of their physical correlates: they would not even be coextensive. Lewis is very much aware of this objection and is concerned to defend his view against it:

> Putnam argues that the brain-state hypothesis (and with it, the functionally specified brain-state hypothesis) ought to be rejected as scientifically implausible. He imagines the brain-state theorist to claim that all organisms in pain--be they men, mollusks, Martians, machines, or what have you--are in some single common non-disjunctive physical-chemical brain-state. Given the diversity of organisms, that claim is incredible. Put the brain-state theorist who makes it is a straw man. A reasonable brain-state theorist would anticipate that pain might well be one brain state in the case of men, and some other brain (or non-brain) state in the case of mollusks. It might even be one brain state in the case of Putnam, another in the case of Lewis. No mystery: that is just like saying that the winning number is 16 in the case of this week's lottery, 137 in the case of last week's. The seeming contradiction (one thing identical to two things) vanishes once we notice the tacit relativity to context of one term of the identities. Of course, no one says that the concept of pain is different in the case of different organisms (or that the concept of winning number is different in the case of different lotteries), it is the fixed concept expressed by 'pain' that determines how the denotation varies with the nature of the organism in question. Moral: the brain-state theorist cannot afford the old prejudice that the name of a necessary being (such as a state) must name it necessarily and independently of context.[14]

This argument is particularly noteworthy in its departure from the typical central state materialist analysis

---

[14]"Review of Art, Mind, and Religion," The Journal of Philosophy 66 (October 1969): 22-27.

(as exemplified by Armstrong and others of the "Australian School"). They require that for every type of psychological state, there be one type of physical state which, in every case, is identical with the corresponding psychological state--i.e., "type/type" physicalism--which is typically contrasted with "token/token" physicalism (as espoused by Fodor and others). For such type/type materialists, an admission such as the above would be tantamount to concession.

So we must apparently conclude that there are significant differences between Lewis and the standard "Australian" line. Still, the above quote creates a tension with Lewis' purported claim that the universal pain is to be identified with 'something that befalls many people at many times'; and with Lewis' categorization of himself as a type/type materialist in the Armstrong mold.[15]

Given our above quote, can we attribute a form of type/type materialism to Lewis (even if it is not precisely the same as the standard version--so long as it preserves the basic spirit)? This is the question which the rest of the section is concerned with. Notice that I am not questioning here Lewis' credentials as a materialist tout court; merely his claim to be a type/type materialist.

---

[15]Thus, "D. M. Armstrong and I (independently) proposed a materialist theory of mind that joins claims of type/type identity . . . ," in "Mad Pain and Martian Pain" forthcoming in Philosophy of Psychology, N. Block, ed.

It is, by the way, only construed as such that Lewis would be the target of the objections of, e.g., Putnam and Fodor.

Have we here been stalking a straw horse? Since after all, Lewis takes mental terms to be analytically equivalent to functional descriptions, should we not place him in the functionalist camp and not the physicalist? One might even think that this business of relativity to context follows from the functional definability of mental terms as understood by Lewis. For the idea is just that such terms refer to whatever happens to fill the functional role: and just as, as we shall see in the next section, such descriptions can refer to different things in different possible worlds, so, too, in the actual world, can they pick out different things in different people.[16]

The answer to our question is: No. We have not been stalking a straw horse. It is an important component of Lewis' views that he is a physicalist and not a functionalist, as is, say, Putnam. For Putnam, pain is (identical to) a functional state and not a physical one; and, consequently, distinct pains do not necessarily have anything physical in common. For Lewis, on the other hand, mental states are (honest to goodness) physical states; they just happen to be functionally specified. To say they are so specified is

---

[16]Lewis himself, in a forthcoming paper, "Mad Pain and Martian Pain," draws the analogy. "If a nonrigid concept or name applies to different states in different possible cases, it should be no surprise if it also applies to different states in different actual cases. Nonrigidity is to logical space as other relativities are to ordinary space.

a comment not on themselves but on our method of referring to them. And this is just how Lewis wants it, for he wishes to show that, 'every experience is some physical state'. So our queries as to the nature of Lewis' physicalism are not misplaced. Lewis is clearly a metaphysical physicalist.

There are two ways of reading the relativization to context quoted just before, which Lewis uses to disarm objections. The first reading finds its greatest support in Lewis' most recent article, "Mad Pain and Martian Pain." On this view, the relativization to context does not manifest itself within the purported definition of 'pain'. Rather, 'pain' is defined simply, and independently of context, as the satisfier of a certain causal role. This definition will, within a given context or "population," pick our some state. For humans, it picks out pain for humans, while for Martians, it picks out pain for Martians, and so on. The one nonrelativized concept--'pain'--picks out <u>different</u> states in different contexts, just as the concept 'the winning number' picks out different numbers in different contexts.[17]

---

[17]Actually, the example of 'the winning number' (or 'this week's winning number') is not a good one. This latter term is an indexical--i.e., a term whose reference varies with the context of utterance: one thing when uttered this week; another when uttered next week. But 'pain' and its ilk, as presently analyzed by Lewis, is not at all like a typical indexical: 'Pain' refers to the same thing(s) in every context of utterance--i.e., different speakers, different times, different audiences, different physical

Thus Lewis says:

> We may say some state occupies a causal role for a population . . . . Human pain is the state that occupies the role of pain for humans. Martian pain is the state that occupies the same role for Martians.
>
> A state occupies a causal role for a population, and the concept of the occupant of that role applies to it, if and only if, with few exceptions, whenever a member of that population is in that state, his being in that state has the sort of causes and effects . . . .

How, on this view, is it assured that there will be one unique physical state corresponding to each population, species say? That, on the present reading, is taken care of by the second premise. It is, in this case, the empirical hypothesis that there will be species specific physical states corresponding, by and large, to each psychological state.

(Fn. 17 continued)

locations, different pointing gestures, etc.; and what it refers to in any such context is different in different entities--again unlike typical indexicals. Thus, in Lewis' writings on semantics, an intension (not precisely a meaning, but close enough for present purposes) is a function not just from possible worlds to things, but from indices. These indices accomodate the contextual elements necessary to determine the extension of indexicals ("General Semantics," in Semantics for Natural Languages, Davidson and Harman, eds.). The contextual elements listed would not, however, help to determine the extension of 'pain'--for the context is irrelevant to its extension.

A better example for Lewis' purposes would have been something like 'the largest city in the state'; a different city for each state. Though even here the analogy is not exact--there is an implicit indexical element missing from the case of 'pain'--we assume that the city in the state in which the utterance occurred, is being referred to.

Given this view, we would have to reformulate the argument at the beginning of the paper by a series of arguments, a schema for which is as follows:

(1) (In all populations (contexts)) Mental State M = the occupant of causal role R (by the definition of M)
(2) (In population A) Neural state N = the occupant of causal role R (by the physiological theory)

---

(3) (In population A) Mental state M = neural state N (by, inter alia, the transitivity of identity).

So 'pain' may pick out different states in different species; indeed, according to our first quote, it may pick out different states in different individuals: one state for Putnam and another for Lewis. What is true is that the concept and the name 'pain' apply equally well to both.

After some reflective thought, this view seems intuitively unappealing. One would have thought that not only is it true that the same concept, 'pain', applies to Putnam and Lewis' states (as the concept 'the winning number' applies to 16 and 137), but that the universal state that 'pain' refers to in both was the same. That both are in the same psychological state--pain--not merely in different states picked out by the same concept.

Now it is true that according to Lewis, there is a state which both Lewis and Putnam are in: the state of having pain. That state, as we shall see in the next section, is a "functional" state distinct from pain. But as we shall also see in the next section, this state is, properly, not taken by Lewis to be physical. So the state

which is had by both Putnam and Lewis is not physical, while states which are physical are not the same for both Lewis and Putnam. Yet, my reflective intuition at least is, that there is one relevant psychological state here; and that is the state which a type/type physicalist must show to be physical (again, above and beyond just showing that distinct states are picked out by the same concept). This, Lewis does not do.

There is a closely related point to the above which, instead of appealing to intuitions, concerns the motivation behind a type/type view. That motivation, as I understand it, is to give an answer to the question of what it is that all pains have in common in virtue of which they are all pains. This is the real issue, as I understand it, that divides the type/type from the token/token physicalists. The latter will claim that though every mental state is indeed a physical state, what all pains have in common in virtue of which they are pains, is <u>not</u> anything physical (such a physicalist might, e.g., say that it is something functional). While, for a type/type physicalist, it <u>will</u> be something physical.

From this perspective, it appears that Lewis' present view is closer to the token/token view; for pain in different species or individuals may have nothing (of relevance) physically in common.

The force of this point is mitigated somewhat if we

simply say that it is part of Lewis' (empirical) claim that the mental states of all members of a species (or most), in fact, have the same physical basis; that if this turns out to not be the case, then his claims for type/type identity will have been falsified. Lewis could then say that there exists something physical in virtue of which all (or most) pains within a species are pain. A similar response would be much more difficult to accept if Lewis is willing to relativize, as seems to be the case, to individuals.

In any case, Lewis' argument, as presently construed, does depend on the empirical assumption that there will be natural kinds, species say, which have specific physiological bases for their mental states. And, of course, the obvious problem with this assumption is that it may turn out to be wrong; though the present evidence does point in that direction.

Even if we grant that Lewis is correct in this assumption, there remain two further points. First, we may construe Lewis' arguments to be concerned strictly with type/type identity theory: the claim that each mental universal is identical to some physical universal. But even if his argument suffices for this claim, it will not do if we take into account the universals of all the other (than everyday psychology) special sciences. For to these, there would seem to be no corresponding physical natural kinds--e.g., money in economics.

Of course, if we restrict our attention to the identity theory, narrowly construed, other special sciences are not relevant. But in point of fact, Lewis' claims for the identity theory are just a part of his larger claims regarding all the sciences. Thus, Lewis' defense of his second premise in, "An Argument for the Identity Theory," reads:

> My second premise is the plausible hypothesis that there is some universal body of scientific theories, of the sort we now accept, which together provide a true and exhaustive account of all physical phenomena (i.e., all phenomena describable in physical terms). They are unified in that they are cumulative: the theory governing any physical phenomenon is explained by theories governing phenomena out of which that phenomenon is composed and by the way it is composed out of them.

It should be clear that Lewis' claims for the identity theory are part of his larger physicalist doctrines. And this is as it should be. One of the reasons that there has traditionally been so much discussion of the identity theory taken as a thesis about mental entities is that it has constituted the greatest challenge to the materialist. Consequently, a view which adequately defended its materialist claims with respect to mental entities, but left itself open to objections with respect to other sciences, could hardly constitute an adequate defense of materialism tout court. And this, I claim, appears to be the present status of Lewis' type/type materialist claims. For there does not seem to be any reason to believe that there will be natural kinds corresponding to the universal

terms of all the special sciences.

If, on the other hand, I am wrong and Lewis does intend his type/type claims to apply to psychology only (by the way, I do not think this is his view), then the claim reduces to no more than an interesting, and probably true, empirical conjecture. But, it can not be said to even purport to be a defense of type/type materialism as a general metaphysical doctrine. Let me point out yet one more time that nothing I say here is meant to impugn any claims about token/token physicalism.

Finally, whether we construe Lewis' thesis as applying to psychology only, or to all sciences, it does seem strange that his claim for (type/type) materialism hinges on a falsifiable empirical claim. For the doctrine is a metaphysical one and not the sort one expects to be open to such invalidation. One would have thought that Lewis would wish to secure his views against such empirical invalidation.

Thus, imagine a possible world in which, for evolutionary reasons, the brains of sentient beings are adaptable to the extent of having their physical composition change over time (perhaps with changes in the environment). Such a world would not be, merely because of the above-mentioned condition, any less a materialist world than our own. But on our present reading of Lewis, he would have to say that such a world is one in which the claims of type/type materialism were not true. This example highlights the

empirical character of Lewis' views as we are presently construing them. The thesis of materialism (type/type or otherwise) is a doctrine about one of the fundamental aspects of the universe; it should not be affected by the vagaries of, say, evolutionary development.

The second reading of Lewis reflects a concern for this last problem and tries to disarm it. On it, Lewis is concerned to defend his claim against the possibility of unfavorable empirical discoveries. This is done by having the concept of 'pain' itself (and of course, other mental terms) be a relativized concept; most perspicuously given as 'pain for Y', the free variable being implicit in our everyday usage of the term 'pain'; and having a value which varies over species, races, individuals, or whatever.

There are several reasons for thinking that something like this is what Lewis has in mind. First, he seems to genuinely desire to protect his view from the vagaries of empirical findings. Thus, consider the following quotes: " . . . pain might well be one brain state in the case of men, and some other brain (or non-brain) state in the case of mollusks. It might even be one brain state in the case of Putnam, another in the case of Lewis." Or: "In general, or in the case of a given species, or in the case of a given person. It might turn out that the causal roles definitive of mental states are occupied by different neural (or other) states in different organisms." Finally, consider this

(from an unpublished thesis):

> John Perry has made an objection to Lewis that is similar to the Putnam objection discussed above. Imagine that there is a radical change in the world at time t_5. Before t_5, the state occupying R in Jones is B (where $A \neq B$). According to the revised definition (2), pain for Jones = A, and pain for Jones = B, which again leads to a contradiction.
>
> Lewis replies to Perry by stating that pain is relative not only to individuals, but also to times. That is, the context dependence of 'pain' extends to times in addition to individuals. If we again revise the definition of 'pain' to account for time, we should define 'pain for (or in) y at t' where 'y' ranges over individuals and 't' ranges over times. Thus, the new definition is:
>
> (3) Def: pain for y at t = the state occupying causal R in y at t.
>
> Definition (3) enables Lewis to avoid the contradiction we got using definition (2). For now we get: 'pain for Jones at t_1 = A and pain for Jones at t_6 = B', which involves no contradiction.[18]

The above quote is admittedly third-hand and can not be relied on as to details. But unless the story is completely apocryphal, it does buttress this second interpretation. For apparently Lewis is concerned to protect his view against the possibility of empirical disconfirmation; even where that possibility is in all likelihood not actual. He feels the possibility needs to be accomodated by his position.

As a further point, notice that, as we quoted before, Lewis says that, "it is the fixed concept expressed by

---

[18]This is from an unpublished thesis by N. Lubow, "The Mind-Body Identity Theory."

'pain' that determines how the denotation varies . . . ." Now for Lewis, a concept (as we shall discuss in detail later) is a function from possible worlds to particular entities in those worlds; in the case of the concept 'pain', those entities would be states. But actually the situation is somewhat more complex: the functions are not from possible worlds merely, but from n-tuples, of which possible worlds are just one coordinate.[19] The other elements are contextual elements: time, place, speaker, etc. In the case of mental terms, it appears that there will have to be one further coordinate--Y let us call it--picking out a particular state based on a given population, however we construe populations.

So that once we have in fact fixed the interpretation of all the other coordinates of the function for a given argument (including possible world, let us say) in the case of the concept 'pain', what we will have will be a function which picks out for each population Y, the satisfier of the causal role for that population; or to formalize, a function which for each value of Y satisfies; $\exists\theta((x)(x \blacktriangleleft Y \rightarrow Q(\theta)))$ where Q represents our postulate (i.e., the conjunction of all our platitudenous mental beliefs). But this concept is just what I have suggested for Lewis on our second reading.

---

[19] See, "General Semantics," in Semantics for Natural Language, Davidson and Harman, eds.

Thus, it is that the relativization to context, on this view, is secured by the very meaning of the terms in question, and is not a question of matter of fact. This interpretation of Lewis, it must be admitted, is not forced on us; but it does appear, the above considerations adequately show I believe, to be a plausible reading of Lewis. And even if it is not Lewis' intention, it is certainly a line, the feasibility of which is worth examining. We shall see later that H. Field's reading of Lewis, in a highly praised article, is very similar.

What sorts of things are we to take Y as ranging over? On this second interpretation, it would not do to say species, or even individuals (as is apparently suggested in the "Mad Pain and Martian Pain" article).

For Lewis, analysis is supposed to capture the meanings of our everyday psychological terms; but why should people happen to mean in their use of everyday mental terms such as 'pain', 'pain for a species'? And further, even if they did, we are granting that it is possible that different individuals within a species may have different psychological states corresponding to 'pain'. Our everyday concept would now have to be narrowed to 'pain for a person'. But once again, why, after all, should, in such an eventuality, people happen to mean in their use of the everyday mental term 'pain', 'pain for a person' (what a happy coincidence)? In short, what reason do we have for

believing that our everyday mental terms are going to have just that restriction of domain which empirically turns out to be that for which there is only one corresponding physiological state. Further, if we grant that, e.g., 'pain' is a relative concept, as Lewis suggests, then it seems intuitive that the relevant populations should be a natural kind and probably at least a species, as he himself says.

Thus, on this second reading of Lewis, where the contextual relativity is construed as an implicit part of the term's meaning, Lewis is driven, if he is going to protect against the possibility of empirical falsification, to something very much similar to what was quoted above: we need to define pain for an individual at a time, or pain for O at t.

But now his claim will be that for any given person at a particular time at a particular place, an occurrence of pain will be some one physiological state or other; a thesis hardly to be distinguished from token/token identity (though we know that Lewis is out to claim more for his view).

There are, to be sure, ways of circumventing this conclusion.[20] Thus, one might say that the view is that in

---

[20] One possibility would be to say that what is identified is a universal with only one instance. Besides being an ad hoc and uninteresting way of saving type/type physicalism, it is certainly something Lewis would not hold. He is concerned with something that occurs to "many people at many times."

any given context (i.e., a person at a place at a time), pain is identified with some one universal physical state. Though from some perspectives, one might discern a metaphysical distinction between such a view and token/token identity, it does not generate an interesting distinction as far as physicalism goes: it does not capture what the type/type theorists were after, and is something that a token/token theorist would find extremely comfortable; for as mentioned previously, the motivation behind the type/type views is to elucidate what all pains (of a certain sort) have in common in virtue of which they are pains (of that sort). And the view we are discussing here is certainly not one which can respond, "something physical"; for each occasion of pain, though it may be something physical, may also be distinct from all others.

To be sure, even on the present view, it is still possible that pain in O_1 at t_1 will be identical with pain in O_2 at T_2 given the appropriate physical circumstances. This does not make the present view any less a token/token view: for similar comments can be made of any token view.

Let me conclude my discussion of this second reading by giving Field's interpretation of Lewis, which is, in many respects, similar to that just given. The relevant passage is the following:

Suppose that ϕ is a psychological theory that is intended to apply any time t to all organisms which are of type Σ at t. (ϕ might, for instance, be a theory intended to apply to all adult humans, or to all organisms capable of feeling pain, or to all rational beings.) For simplicity, let us suppose that ϕ is finitely axiomatized; then we can represent it as a single formula which I abbreviate as $A(\chi,t)$, where χ is a variable ranging over organisms and t a variable ranging over times. [Then the theory is true of all the organisms in its intended range if and only if the following claim holds:

> For any t and any χ, if χ is of type Σ at t then $A(\chi,t)$.

It is this last claim, rather than ϕ, that is properly speaking true or false, so you might prefer to use the term 'theory' for the last claim rather than for ϕ.] In giving crude formulations of psychological theories we often omit the variables (and initial qualifiers), but they must be understood as implicit: we say 'pain has such and such a causal role' when what we really mean is 'For any t and any χ of type Σ at t, pain has such and such a causal role in χ at t'. If we do not write the theory in this way, we cannot properly define the notion of a realization.

Suppose that the specifically psychological primitives in ϕ are $T_1, \ldots, T_n$; then we can write ϕ as $A(T_1, \ldots, T_n; \chi, t)$. For simplicity, I will assume that $T_1, \ldots, T_n$ are all predicates. Let us say that an n-tuple $\langle P_1, \ldots, P_n \rangle$ of properties realizes ϕ in organism X at time t if and only if the formula $A(Y_1, \ldots, Y_n; \chi, t)$ is true of $\langle P_1, \ldots, P_n, X, t \rangle$; and that such an n-tuple uniquely realizes ϕ in X at t if it and no other n-tuple realizes ϕ in X at t.

Now if ϕ is a psychological theory with n primitive psychological predicates we can use it to define n functional properties. Suppose for instance that T_j is a predicate that stands for a 1-place property of organisms, the kind of property (like pain) of which it makes sense to say that the organism has it at one time but not at another. (Lewis calls properties of this sort 'states', but I prefer to reserve this term for a different use.) If T_j is a predicate of this sort, then the jth functional property associated with ϕ is the property ψ defined as follows:

(2) X has ψ at t $\equiv_{df}$ there is some 1-place physical property P such that

(i) P is the jth component of a unique realization of ϕ in X at t

(ii) X has P at t

If ψ is the jth functional property associated with ϕ, we can then say that a realization of ψ in X at t is simply the jth component of a unique realization of ϕ in X at t. From this and (2) we derive

(2') X has ψ at t if and only if there is some 1-place physical property P such that

(i) P realizes ψ in X at t, and

(ii) X has P at t.

This machinery enables us to give a precise sense to the general remarks of five paragraphs back. What functionalism about pain claims is that the property of pain is a functional property associated with some theory ϕ, by (2)(or by the analog of (2) with the word 'physical' replaced by 'non-functional'--see note 26). By taking functionalism in this way, we can make precise sense of various vague notions appealed to in the general remarks (e.g., the notion of psychological isomorphism), and we can also verify the really important claim that if materialism is true, then for an organism to have the psychological property of pain it must have some physical property that realizes that psychological property in the organism.[21]

As to the correctness of this interpretation, Field says:

> Lewis' own account of what realizations are is strictly speaking inaccurate because he does not take the precaution I've recommended: as his account stands, something can serve as a realization of pain only if it realizes pain in all organism at all times, thus depriving functionalism of its point. (cf. Harman [11], Ch. 3, Section 4). But it is clear that what I have suggested (which is equivalent to what Harman suggests) is what Lewis really had in mind.

Let us note a few things. 'Pain' as defined here, is a functional property; what he has defined is 'x's having ψ at t.' What this definition entails for materialism

---

[21]This and all the following quotes are from H. Field, "Mental Representation," Erkentniss 13 (1978): 9-61.

is that, "for an organism to have the psychological property of pain, it must have some physical property that realizes the psychological property in the organism" (emphasis mine); i.e., that there is some property in each case, but perhaps different in different cases, which realizes (not, is identical to) the psychological property. But this is precisely token/token materialism. And this is all that Lewis' definitions, as Field and I have construed them, will yield.

It is true that what Field defines as pain is the functional property which Lewis takes to be the property of having pain. Thus:

> According to Lewis, 'the property of pain' and 'pain' refer to different properties: 'the property of pain' refers to a functional property, and 'pain' refers (in the context of discussing a specific organism X at a specific time t) to the non-functional property which realizes the functional property in X at t. I have tried to remain neutral on the question of whether 'pain' refers to the functional property or refers (in a context-dependent way) to a realization if it.

If we do take it in the latter way--that is, if we do construe 'pain' as the non-functional property--then x and t are most appropriately taken to be open variables, or schematic letters. This would essentially yield the prior formulation. In any case, the important point is that we are relativizing to a person at a time. And that is, in spirit at least, a token/token view. Thus, our overall conclusion for this second reading of Lewis: its spirit and possibly even its letter, is that of a token/token view.

As my last major point, I turn to the question of how Lewis' relativization to context affects the subsumtion of everyday mental definition to the more general case of theoretical definition. This is of considerable concern, for, as indicated at the beginning of this section, much of the persuasiveness of Lewis' definitions of mental terms comes from this subsumtion.[22]

This point is applicable mainly to the second reading of Lewis, but also to a certain extent, the first. On the second interpretation, Lewis relativizes to a person at a time. And even on the first reading, it still seems to be true that Lewis is willing to restrict his empirical claim as far as he needs; i.e., he is willing to go with the claim that pain picks out a different universal state in different individuals, if need be.

Now this in a sense trivializes Lewis' claim to have defended a type/type view. Thus, consider the consequences of applying this relativization context to all theoretical terms.[23] This would essentially make any theory immune

---

[22]Let me remind the reader that without it, we must worry about the plausibility of Lewis' definitions (independently of the issue of relativity to context). We then would have to take into account such worries as phenomenological qualities (see, e.g., Block, "Are Absent Qualia Impossible," PR, April 1980). We would also have to worry about "Nagel's Problem," as mentioned in the text.

[23]Of course, in a sense, scientific theories do typically relativize; their laws are meant to apply to a specific domain only. But in such cases, there are independent theoretical

from falsification. Thus, consider if there were some theory positing some specific causal agent--Tox--as the cause of toxemia (a disease of pregnancy). In point of fact though, it turns out that there is no specific cause of the disease, there is only a set of symptoms (or perhaps there are two factors, each of which is sufficient to cause the disease). We could still preserve the truth of our theory, and our purported reduction of the universal, Tox, to a physical entity, by relativizing as appropriate: to individual pregnant women, perhaps. But to be able to do this would trivialize both any purported claims of reduction of, Tox, the relevant universal, and the claim of falsifiability for our theory of toxemia.

Thus, either (1) the relativization move is not extendable to scientific theories generally, including a future scientific psychology, in which case Lewis' everyday definitions lose the virtues of being subsumed under the more general case; or (2) the move <u>is</u> extendable, in which case Lewis' construal of theoretical definition and reduction is vitiated.

Let me here summarize the major points made so far. Two interpretations were given of Lewis' attempt to preserve a form of type/type physicalism. They were seen to have complementary faults. On the first interpretation, Lewis'

(Fn. 23 continued)

reasons for the domain restrictions. In the present case, it is the ability to relativize as one sees fit (i.e., to save one's theory or reduction) that is at issue.

view can legitimately be claimed to be a type/type one, but its domain is restricted, and its truth is contingent on the existence of certain empirical facts, facts which it is physically possible do not obtain. While the second interpretation does have unrestricted domain, and does attempt to insure itself against empirical fansifiability, it can not, however, be said to properly be a type/type view. And both views seem to run into trouble when we consider extending relativization to the special sciences.

Lewis' problems, it seems to me, arise from a set of desiderata which together entail the denial of a truth. He wishes to contend that there are true special sciences (among which we may include folk psychology), that the theoretical terms therein uniquely refer, and that--in line with his reductionist beliefs--all such terms, including those for universals, will ultimately be seen to refer to physical entities. These are, however, jointly inconsistent with the claim, which appears true, that there are true special sciences with more than one physical realization.[24]

---

[24]Putnam has made this point in various contexts, and with a considerable amount of force. For a concise statement, see J. A. Fodor: "Special Sciences (Or: The Disunity of Science as a Working Hypothesis)," Synthese 28, 1974.

Of course, Lewis claims that it is reasonable to expect that theories will have only one realization. Indeed, his definition of theoretical terms embodies that expectation--given his meanings of theoretical terms the theories which contain them logically imply that the theory is true only if it has a unique realization.

Before proceeding to the next section, we must consider one other possible response which, though he does not make it,[24] is open to Lewis; that of disjunctive properties. 'Pain' and other psychological terms, it might be claimed, refer to certain disjunctive physical states: e.g., pain in mollusks v pain in humans v..... ; that is the state one is in when one is in either pain in mollusks or pain in humans or in .... etc.

The problems with such a view have been discussed in various places (e.g., by Block and Fodor in, "What Psychological States are Not," and in Davidson: "Mental Events"). Many of these objections have had, as their crucial premise, that even if a psychological predicate is coextensive with a certain (disjunctive) physical predicate, it still remains to be shown that the coextensivity is lawlike. I am not convinced of the strength of this objection (it strikes me as question begging), but in any case, it is not clear how it is directly applicable to Lewis'

(Fn. 24 continued)

If we grant the reasonableness of this last expectation, then, it seems to me, we can not grant that there is a unique physical realization of such theories. For surely, economics, linguistics, psychology, and even biology, can be satisfied by various physical systems. If we wish then to preserve uniqueness, we will have to allow unique nonphysical universals having various physical instantiations.

[25]Lewis eschews such a move in, "Mad Pain and Martian Pain," Op. Cit.

views.[26]

I can think of three problems for Lewis besides its ad hoc nature, with such a view. First, it is not compatible with what he says--i.e., with construing 'pain' as analogous with 'the winning number'. We could not say that 'the winning number' refers to some one disjunctive number: 16 v 137 v.....

Secondly, imagine that in the actual world, P_1, P_2, P_3 are the physical correlates of pain. On the present view, pain would be identified with the disjunctive state $P_1vP_2vP_3$, as the unique satisfier of the appropriate causal role. However, imagine that on some other world, there is an additional correlate to pain--P_4--which is not correlated with pain in this world; on that world, pain is to be identified with $P_1vP_2vP_3vP_4$. Clearly, this latter disjunctive state would also satisfy the appropriate causal role in the

---

[26]This objection usually comes down to the claim that the two predicates are not necessarily coextensive. While Lewis might grant this, he would claim that he has an argument to show that the referent of 'P-ness'--the name of a psychological state--is identical with the referent of 'Q-ness', a certain physical name. If this is so, then naturally, what the name 'P-ness' refers to in the actual world is coextensive, in all possible worlds, with what 'Q-ness' refers to in the actual world. There is one property (state) referred to by two names, and it has whatever extension it has (naturally) in all possible worlds. It is true, as our intuitions might inform us, that P-ness might not have been Q-ness, but since 'Q-ness' is, on Lewis' view, a nonrigid designator, this is no stranger than it being the case that Benjamin Franklin might not have been the inventory of bifocals. These issues will be discussed in greater detail in the next section.

actual world; thus, if an organism has its skin cut, it is usually in $P_1 \vee P_2 \vee P_3 \vee P_4$; that is, it is either in P_1 or P_2 or P_3 or P_4, since by hypothesis, it is either in P_1 or P_2 or P_3. So now there are at least two states which satisfy the appropriate causal role. So the move to disjunction yields multiple realization.

I realize that this is a contrived counterexample but I am not moved: the whole idea of adverting to disjunctive states is contrived to begin with. Such disjunctive states (or rather predicates) are not, in Goodman's terminology, projectible. My objection to the proposal is in the same spirit in which it was offered.

Thirdly, and most importantly, even if we neglect the above two problems and assume disjunctive properties do the trick, allowing these properties is either to be restricted to everyday psychology, or not, and is to be allowed to apply to the referents of all theoretical terms. Either of these alternatives, as with the relativization to context move, will prove to be unsatisfactory.

SECTION II

II

This section shall deal with issues related to Lewis' physicalism arising from his views on semantics and other possible worlds. I shall ignore here the problems raised in the previous section and shall assume, for expository reasons, that pain (and likewise, for other psychological states) is identical with some one physiological state--in particular, with C-fiber stimulation (cfs).

As an entry point into these issues, consider the following two apparent problems for Lewis. First, there is Kripke's argument that if a is identical with b, then necessarily, a is identical with b.[1] It is not the case, however, that necessarily, pain is identical with C-fiber stimulation; for--as Lewis readily admits--in some possible world, pain might be (say) D-fiber stimulation. So if Kripke's argument is sound, it appears that pain and C-fiber stimulation could not be the same, even in the actual world.

Secondly, in Lewis' view, a property is uniquely determined by which things have it, in every possible world; properties a and b are identical if they are coextensive in every possible world. But pain and C-fiber stimulation are not coextensive in every possible world and hence must be distinct.

---

[1]Or, to put it in the formal mode, if a and b are rigid designators, then if 'a=b' is true, so is 'necessarily a=b'.

Lewis is not perturbed by these arguments (he deals directly only with the second, but it is clear how he would respond to Kripke). To understand his response, it is necessary to examine in some detail the general view which underlies it. In a general discussion of theoretical terms, he says:

> I take it that a property is identified when, and only when, we have specified exactly which things have it in every possible world. And I take it that a name of the form 'the property of doing so-and-so' names the property that belongs, in the world w, to whatever does so-and-so. For instance, 'the property of having t_1' names the property that belongs, in any world w, to exactly those things which, in the world w, have the property named by t_1.
>
> Now we can see the problem. Do we mean:
> 1) the property that belongs, in any world w, to exactly those things which, in the world w, have the property named by t_1 in our actual world? That, of course, is just the same property which is named by t_1 in our actual world. On this first reading, 'the property of having t_1' and 't_1' do both name the same property.
>
> Or do we mean: 2) the property that belongs, in any world w, to exactly those things which, in the world w, have the property named by t_1 in the world w? On this second--and, I believe, better--reading, 'the property of having t_1' is a logically determinate name of a certain property, which we may call the diagonalized sense of t_1. The sense of t_1 may be represented by a function $||t_1||$ which assigns to any world w a property $||t_1||$ w. A property in turn may be represented by a function P which assigns to any world w the set Pw of things which, in the world w, have the property. Then the diagonalized sense of t_1 is the property whose representing function assigns to any world the set of things $(||t_1||\ w)w$. It is not named by t_1 in any world, unless T is a very peculiar theory. Neither is it the sense of t_1; that is, not a property at all, but rather a function from worlds to properties.[2]

---

[1] "How to Define Theoretical Terms," Op. Cit. Pg. 437.

Notice that if 't_1' is a rigid designator,[3] there is no distinction between t_1 and the property of having t_1--even on the proprietary reading. The representing function of the property of having t_1 assigns to each world w the set of objects that have the property that the sense of 't_1' assigns to that world--and that is the same property in every possible world. Thus, the sense of 'tallness' (assuming it is rigid) assigns to every possible world the same property (tallness). Consequently, the property of having tallness will have, as its extension, in every possible world, the extension of the same property--tallness. Tallness, and the property of having tallness, are, therefore, the same property.

So, if the distinction between t_1 and having t_1 is to be made to do any work in our present consideration, it must be that 'pain' is a non-rigid designator--and that is precisely what Lewis asserts:

> We must not identify an experience itself with the attribute that is predicated of somebody by saying that he is having that experience. The former is whatever state it is that occupies a certain definitive causal role; the latter is the attribute of being in whatever state it is that occupies the causal role. By this distinction, we can answer the objection that since experience-ascriptions and neural state ascriptions are admittedly never synonymous, and since attributes are identical just in case they are predicated by synonymous expressions, therefore, experiences and neural states cannot be identical attributes. The objection does establish a nonidentity,

---

[3]Lewis does not use the term 'rigid designator'. He uses the term 'non-contingent name', which means more-or-less the same thing. Since the former term has become so popular, I shall consistently use it, even when talking about Lewis' views.

but not between experiences and neural states. (It is unfair to blame the identity theory for needing the protection of so suspiciously subtle a distinction, for a parallel distinction is needed elsewhere. Blue is, for instance, the color of my socks, but blue is not the attribute predicated of things by saying that they are the color of my socks, since ' . . . is blue' and ' . . . is the color of my socks' are not synonymous.)

He develops this point further in an appended footnote:

> Here I mean to deny all identities of the form 'α is identical with the attribute of having α', where α is an experience-name definable as naming the occupant of a specified causal role. I deny, for instance, that pain is identical with the attribute of having pain. In my theory, 'pain' is a contingent name--that is, a name with different denotations in different possible worlds--since in any world, 'pain' names whatever state happens in that world to occupy the causal role definitive of pain. If state X occupies that role in world V while another state Y (incompatible with X) occupies that role in world W, then 'pain' names X in V and Y in W. I take 'the attribute of having pain' on the other hand, as a non-contingent name for the state or attribute Z that belongs, in any world, to whatever things have pain in that world--that is, to whatever things have in that world the state named in that world by 'pain'. (I take states to be attributes of a special kind: attributes of things at times.) Thus, Z belongs to whatever things have Y in W; hence, Z is identical neither with X nor with Y.
>
> Richard Montague, in "On the Nature of Certain Philosophical Entities", Monist 53 (1969): 172-173, objects that I seem to be denying a logical truth having as its instances all identities of the form 'α is identical with the attribute of having α', where α is a non-contingent name of a state which is (either contingently or necessarily) an experience. I would agree that such identities are logically true; but those are not the identities that I mean to

> deny; since I claim that our ordinary experience-names--'pain' and the like--are contingent names of states.[4]

The above quotes are sufficient to respond to the objections raised; some elucidation is, however, called for. First, let me dispell a possible confusion. There is nothing special about 'pain' that makes it nonrigid; most theoretical terms defined à la Lewis will, it appears, be nonrigid.

Let me also emphasize the difference between names which refer to properties and predicates which express or attribute those properties. The sense of a predicate is the property which it expresses--i.e., a function from possible worlds to individuals. The sense of a name of such a property (e.g., 'cleverness') is a very different thing. It is a function which assigns to every possible world a unique property, each one of which is, as above, also a function from possible worlds.

Thus, if one were to ask the average academician on the street, 'Is having the property I am thinking of necessarily coextensive with cleverness?', he would very likely respond, 'No'. Lewis would respond: If you mean to be talking of what the two terms 'cleverness' and 'the

---

[4]"An Argument for the Identity Theory," Pp. 164-165. The problem Lewis is here replying to is not precisely the one I mentioned. The problem here is that there seems to be two non-synonymous predicates expressing the same property--which Lewis believes can not be. Lewis' reply would, however, be the same for either problem; in fact, in Lewis' view (we may take it), the sense of a predicate is the function which assigns to every possible world the extension of the predicate in that world--in which case we really do have one problem here (see below).

property I am thinking of' refer to, then they refer to the same property, which is, of course, necessarily coextensive with itself. If you mean, instead, the sense of the terms 'cleverness' and ' . . . has the property I am thinking of', then you are comparing apples to oranges: the former assigns a property to each possible world, while the latter assigns a set of objects (its extension). Only if you compared the reference of 'cleverness' with the sense of ' . . . has the property I am thinking of' (or equivalently the reference of 'The property of having the property I am thinking of') would you get a meaningful no.

When forming one's theoretical and psychological definitions, according to Lewis, one first reformulates the postulate in such a way that all theoretical (psychological) predicates are eliminated in favor of names of properties, states, etc. This is done by having in the O-vocabulary (non-theoretical--i.e., old, already understood) copulas of the form, '____has the property____', '____is in the state ____at time____'. What goes in the appropriate blank and gets replaced by a variable in the modified Ramsey sentence, is the name of a property or state. So it is pain, and not having pain, that Lewis is primarily concerned with.

Pulling together the various strands so far presented (at the risk of some redundancy), and filling in some of the detail, we have: predicates express (or attribute) properties which are the senses of those predicates. In, 'My socks are blue', ' . . . are blue' expresses the property of being blue.

Two predicates express the same property if they are synonymous--since properties are the senses of predicates. These properties (senses) are "various set theoretical constructions out of worlds and individuals . . . there is no reason to believe that besides those constructions, there are some other entities--the properties, states, etc. themselves."[5] States, as well as events, phenomena,

---

[5]Quoted from private correspondence.

The views I am presenting here have been gleaned from Lewis' published writings. In correspondence, he has made it clear that though I have indeed stated his favored view, I have misled about the degree of determinacy he thinks there is to be found. (My excuse is that I had, in any case, to put some one formulation down to sink my teeth into.)

Thus, Lewis says that properties correspond to set theoretical entities. But which ones? There is no determinately correct answer to this question (Lewis believes). There is no fact of the matter; just as there is not (Lewis believes) concerning the question of whether the ordered pair $\langle x,y \rangle$ is $\{\{x,y\},\{x\}\}$ or $\{\{x\},\{x,y\}\}$. I have identified what predicates express--their senses, properties, and functions--from possible worlds to extensions, and have asserted that states are just a species of property. All this also seems to be, in Lewis' view, one possible formulation among many equally correct ones. One thing that is determinately true, however, is that properties are identical if the predicates which express them are synonymous (that is, they have the same sense) if the functions which correspond (or are identical) to these properties are identical. And this contention is sufficient for the purposes of this paper.

In the rest of the paper, I shall take the views presented above to be Lewis'. The points I make will, I think, be reformulable to coincide with other possible formulations of Lewis' views. If any of my points do depend on the particular manner in which Lewis has formulated his claim, then Lewis can extricate himself by choosing some other. In that case, my argument can be construed (more weakly) as showing that either Lewis must give up the view I am attacking (if my argument is sound), or he must give up the contention that it is arbitrary which formulation one uses concerning properties, states, etc.

experiences, etc., are just a species of property: functions from worlds to sets of time slices.

Names refer to, among other things, properties (states, etc.). 'Blue' names a property--that which is expressed by ' . . . is blue'. But the sense of 'Blue' is not identical with the sense (that which is expressed by) of ' . . . is blue'.[6] The sense of 'Blue', as with all names, is a function whose value in each possible world is some one entity which the name refers to in that world--in this case, a property. While the sense of ' . . . is blue' assigns to each possible world sets of, mostly, relatively mundane objects (e.g., socks).

Finally, and of most relevance to us, there are cases in which @ (some property name) does not even name what ' . . . has @' expresses, unlike the case of 'Blue'. This will occur when and only when @ is a nonrigid designator. Thus, 'The property I am thinking of' names, in the actual world but not in every possible world, stupidity say. So what this term actually names is a property whose extension in every possible world is the set of stupid things. ' . . . has the property I am thinking of' expresses (in, say, 'Harry has the property I am thinking of') the property whose extension in any possible world is the set of things I am thinking of in that world: stupid things in this world, brilliant ones in some other. Since their extensions

---

[6]And named by, 'The property of being blue'. Where @ is some predicate, 'The property of being (or having) @' is taken to name what is expressed by ' . . . is @'.

are possibly different, the property I am (actually) thinking of and the property expressed by ' . . . has the property I am thinking of' are therefore distinct; and similarly for any nonrigid designator of a property.

In particular, this situation obtains for Lewis in the case of 'pain'. There are four, not necessarily distinct, things to consider. The state named by 'pain'; the state expressed by ' . . . is in pain' (and named by 'The property of being in pain'); a certain physiological state named by 'cfs'; and what is expressed by 'is in cfs'.

Let us assume that 'cfs' is rigid. It names in every possible world the same state. What it names is then identical, as should be clear from the above, with that which is expressed by ' . . . is in cfs' (and named by 'the property of being in cfs').

'Pain', however, is nonrigid; as it turns out in the real world, it names (we are assuming) cfs. So pain = cfs = the property of being in cfs. But none of these is identical with the property of having (or being in) pain--the property we attribute to people when we say they are in pain. That property, since 'pain' is nonrigid, is the distinct property whose extension in any possible world w_1 is the set of things which are in pain ('pain' can here be replaced by the nonrigid description to which it is synonymous) in w_1.

All of the above points are summarized in the following table. The reader should satisfy himself that he understands and is in agreement with everything in it.

LEWIS ON UNIVERSALS

Universals, properties, attributes, states, events, experiences, kinds, senses, etc., are all set theoretic constructions out of possible worlds and individuals. Some of these may be subsets of others--e.g., the extensions of states are particulars at times. Predicates express or attribute properties which are the predicates senses--viz., the functions which assign to each possible world the set of individuals which satisfy the predicate in that world.

Term	Rigid/Nonrigid	Reference/Extension	Sense
1) 'cleverness' ('cfs')	rigid (let's agree)	Reference: The property, C, which is the sense of ' . . . is clever'	A function //cleverness// which assigns to each possible world w, a property, $//cleverness//_w$; in this case= C for all w
2) ' . . . is clever' (' . . . is cfsing')	Not applicable to predicates	Extension: all those things which are clever	A function (property), A, which assigns to any world w, the set of things C_w which, in the world w, are clever.
3) 'The property of having cleverness (being clever)' ('The property of being in cfs')	rigid	Reference: the diagonalized sense of 'cleverness'; a property which assigns the set of things: $(//cleverness//_w)_w$ to each world w. Identical to C, since for all w, $C_w = (//cleverness//_w)_w$	A function //The property of having cleverness//, which assigns to each world a property. //The property of having cleverness$//_w$; in this case= the diagonalized sense of 'cleverness' = C, for all w

Term	Rigid/Nonrigid	Reference/Extension	Sense
4) 'The property I am thinking of' ('pain')	nonrigid	Reference: The same as that of 'cleverness'. C	A function // The property I am thinking of// which assigns to each world w, a property //The property I am thinking of//$_w$; in this case, different properties for different w
5) ' . . . has the property I am thinking of' ('has (is in) pain') Lewis' preferred analysis	Not applicable to predicates	Extension: all those things which have the property I am thinking of; same as the extension of ' . . . is clever'	The diagonalized sense of 'The property I am thinking of'; a function (property) which assigns to each world w, the set of things (//The property I am thinking of//$_w$)$_w$
6) 'The property of having the property I am thinking of' ('The property of having pain') Lewis' preferred analysis	rigid	Reference: sense of 5)	A function //the property of having the property I am thinking of//, which assigns to each possible world w, a property, //The property of having the property I am thinking of//$_w$, the same property for all w

We can finally take care of the problems with which this section opened. Kripke's claim that pain might not be cfs proves nothing since 'pain' (in Lewis' view) is a nonrigid designator. This example proves no more mysterious than the fact that the inventor of bifocals might not have been Benjamin Franklin.[7]

As for the second problem, pain and cfs do have the same extension in every possible world--they are, after all, the very same state (property). To be sure, in a sense, pain might not have been cfs, for 'pain' does not necessarily refer to the same state as 'cfs': the concepts (the senses of) 'pain' and 'cfs' are distinct. This should not bother us, for in a similar vein, pain might not have been pain: 'pain' does not necessarily refer to what 'pain' actually refers to.

What may be responsible for confusion is, Lewis believes, the failure to distinguish the above claim that pain and cfs are identical, which is true, from the claim that the property attributed to people when we say they are in pain is identical with the property we attribute to people

---

[7] Kripke, of course, thinks that terms like 'pain' are rigid designators; Lewis does not, so there is at least a stalemate until the question of rigidity can be cleared up. Kripke might try to reformulate his argument by saying, 'niap' say, and let us proceed to give the same argument using this term. To this, Lewis might rightly reply that rigid designators are not so easily created. One must give a word meaning and in the present case, 'niap' seems to be synonymous with 'pain' (which is nonrigid). If we clearly did have a rigid designator--e.g., "That entity which, in fact, 'pain' refers to in the actual world," our intuitions about the possibility of the entity referred to not being identical with cfs would no longer be clear (after all, isn't cfs what 'pain' refers to in the actual world?). Kripke has a reply, I believe, but it would take us far afield.

when we say they are having cfs, which is false. And that confusion, thinks Lewis, is due to a failure to distinguish pain from the property of having pain.

Thus, it appears that Lewis has easily circumvented two rather difficult problems. It might, at first glance, be thought that this escape is vitiated by the apparently gratuitous assumption that 'pain' and other psychological terms are nonrigid. 'Pain' and its ilk might, after all, be naturally taken to be on par with 'cow', 'gold', and other natural kind terms; which act like, according to the popular account of Kripke and Putnam, rigid designators.[8]

Though I do indeed think something like this correct, I shant argue for it here. Perhaps ordinary psychological terms are nonrigid. In any case, the important issue (see section 1) is whether the theoretical terms of a mature psychology, if we ever have any, would be reducible to physical terms. Our concern should then be whether <u>such</u> theoretical terms are, as Lewis suggests, nonrigid. And to this question, the Kripke-Putnam analysis is not directly

---

[8]But not, of course, according to Lewis' account.
Even if the Kripke-Putnam analysis is correct, it could turn out that 'pain' does not, in fact, name any natural kind. This is no problem as it is true of most terms. 'Richard Nixon' might not name a person, or anything else for that matter. We feel no inclination on that account to say that it is not a rigid designator; for it <u>purports</u> to designate (rigidly) a certain person. Similarly, it might be argued, we purport by our use of 'cow' to name a certain species, and by our use of 'pain' a certain kind of mental entity, if anything.

germane: their examples are not of theoretical kinds, and there is no saying that their intuition holds up for such.[9]

In sum, Lewis' views deal very neatly with the two problems which opened this section. It is not the least virtue of his account that it does so, especially considering how easily other views are ensnared.

His responses to these problems, as I have conveyed them, bring to the fore two assumptions which call for consideration: The first is whether or not to take, in the manner of Lewis, theoretical terms as typically nonrigid. This question has obvious interest beyond that which concerns the meaning of everyday psychological terms. The second is the legitimacy of Lewis' distinction between pain and the property of having pain, upon which his claim that 'pain' refers to some physical state seems to depend. I shall deal, in reverse order, with these questions in the following sections.

---

[9]In, "Reference and Theoretical Terms," Nous Vl. X, #3, Sept. 1976, Enc has argued that in fact the Kripke-Putnam account is deficient when applied to theoretical terms. Enc's argument would be of no solace to Lewis though; it argues against his account even moreso.

SECTION III

III

A careful reader will have noticed that the previous section's distinction between pain and the property of having pain has the untoward consequence that the property of having pain, as construed by Lewis, is not physical. Here we have yet another important respect in which Lewis' view differs from traditional "Australian" materialism.[1] For the Aussie view--a view one would think an integral part of any materialism--is that all properties, events, etc., are physical properties, events, etc. Further, the property of having pain is precisely the sort of property the Aussies are typically concerned to insist is physical. Thus: "I am not arguing that the afterimage is a brain process, but that the experience of having an afterimage is a brain process."[2]

One might contest what I have just said: The property of having pain being, after all, a set theoretic construct is, thereby, a physical property for Lewis. Let us quickly recall what (as quoted on pg. 35) this property is: Assume 'pain' is synonymous with a description which picks

---

[1] Recall that the other respect was Lewis' allowing different physical states to correspond, in different contexts, to the same mental state.

[2] J. J. C. Smart, "Sensations and Brain Processes," Philosophical Review, LXVII (1959): 141-156. One should, in the above quote, replace 'afterimage' by 'pain'. The view quoted takes care of afterimages (pains) by saying that there is, in a sense, no such thing.

out some physical state, X, in this world, and Y in world w. Then the experience (property, state, attribute) of having pain is the experience (property, state, attribute) whose extension in the actual world is X's extension, and in w is Y's extension; it is, consequently, identical to neither of these physical states; here, by physical, we must mean something like, referred to by a term of physics or a term which can be appropriately reduced. Allowing that the mere fact that a property is a set theoretic construct to purchase its physicality would be buying it very cheap indeed.[3]

Nor should one think that the nonphysicality of having pain is a consequence of the particular conventions used by Lewis--conventions which could be changed if need be. Lewis is categorical concerning the adduced distinction between t and 'the property of having t'. And it is this distinction which is invoked to resolve the problem--a very real problem for Lewis--of the nonsynonymy and non-necessary coextensivity of the properties attributed in 'John is in pain' and 'John is in cfs'. Lewis' distinction is not a mere matter of convention; whatever formulation he chooses must deal with the nonsynonymy problem.

---

[3]H. Field makes this point in, "Mental Representation," Erkentniss 13 (1978): 9-61. I.e., if the above sort of move were legitimate, we could always preserve our materialism by reducing any recalcitrant property to the set of things, in each possible world, which have that property. This would pretty well make one's physicalism vacuous.

Lewis would not, however, be perturbed by our concerns; for he himself says, or at least implies, that having pain is not physical; it is, in any case, part of his general view that there can be nonphysical entities.

But there is more to the problem than we have so far presented; it is not merely that Lewis countenances nonphysical entities; it is the nature of the entities so countenanced. For having pain, or better, the property of being a pain, which would get similar treatment by Lewis, is a property whose physicality is crucial for physicalism.

It exemplifies, on Lewis' analysis, that in virtue of which all pains are pains, what all pains have in common is that they satisfy a certain causal role; the property of satisfying that causal role is the property of being a pain; the property of being in a state which satisfies that role is the property of being in pain. These are precisely, or so it seems to me, the crucial properties for a type/type physicalism to reduce (doing so is a large part of the traditional motivation behind a type/type as opposed to token/token form of materialism).

What we must conclude is that (1) Lewis' view substantially diverges from standard central state materialism, and (2) it is, in certain respects, an attenuated version of physicalism. In the first section, we questioned whether Lewis differed significantly from proponents of a token/token view; here, we are questioning his credentials as a physicalist.

This point can clearly not be pressed too far; Lewis, and others, will accept the consequences with equanimity--and not feel they are any the less a physicalist thereby. Those who are physicalists must decide for themselves whether they could live with this version of it.

Unlike that which we have just discussed, the relation between the problems that follow and the conventions used in Lewis' formalism is not clear. All of them, though some less than others, depend to some extent on these conventions; whether they depend crucially, or would apply no matter what set of conventions Lewis used, is hard to say. If the following arguments are indeed convention dependent, then they should be construed as directed not at Lewis' substantial claims, but at (1) Lewis' present formulation, and (2) the claim that it is arbitrary which set of conventions one uses (See fn. 5, section II).

The first point is a relatively minor one. Recall that in Lewis' formulation, all theoretical (psychological) predicates are replaced by names; this is how we come to talk of 'pain' as distinguished from ' . . . is in pain'. But though 'pain' may look like a name, in practically all of its everyday uses, it does not function as such--e.g., 'I am in pain', 'He has a pain in his foot', etc.[4] Of course, it is Lewis' formulation to do with as he pleases:

---

[4]Thus Davidson remarks in a different context:

> We recognize that there is no singular term referring to a mosquito in 'there is a mosquito in here'

in particular, he can restrict his formulation to names. But in the case of folk psychology, for which he purports to be giving the meaning of our everyday terms, this sort of formulation is embarrassing, to say the least.

The next point is that Lewis' formulation would count as nonidentical properties which most physicalists would claim are. Thus, most would agree that the property of having a certain temperature is identical with the property of having a certain molecular energy. Lewis will not. This is essentially the same point as was made with the property of having pain, but applied now to a less contentious example.

When one has a name for an experience such as 'pain', it may be reasonable to distinguish the experience from

---

(Fn. 4 continued)

> when we realize that the truth of this sentence is not impugned if there are two mosquitos in the room. It would not be appropriate if noticing that there are two mosquitos in the room, I were to ask the person who says, "There is a mosquito in the room," "Which one are you referring to?" In the present analysis, ordinary sentences about events--i.e., 'Doris capsized the canoe yesterday'--are related to particular events in just the same way that 'There is a mosquito in here' is related to particular mosquitos. It is no less true that Doris capsized the canoe yesterday if she capsized it a dozen times than if she capsized it once; nor if she capsized it a dozen times does it make sense to ask, "Which time are you referring to?" As if this were needed to clarify 'Doris capsized the canoe yesterday'. We learned some time ago, and it is a very important lesson, that phrases like 'a mosquito' are not singular terms, and hence do not refer as names or descriptions do . . . . ("The Individuation of Events," Essays in Honor of Carl Hempel, Rescher ED.

Davidson's point can be made with equal validity about the occurrence of 'pain' in most sentencnes; the truth of the sentence is not impugned if there is more than one pain.

the attribute of having that experience. But there are many everyday psychological terms for which this distinction can not be easily made. Thus, suppose that it is claimed that intelligence is identical to I-fiber stimulation (or, e.g., getting correct scores on an intelligence test). The senses of the predicates in 'John is intelligent', and 'John is having IFS', are not synonymous; the question now is, can Lewis deal with this fact in a manner analogous to his treatment of the corresponding problem in the case of pain?

I think not. Consider: if we did make such a move, we would have to say that there is a property attributed to people when we say they are intelligent (call it what you will: the property of being intelligent or the property of having intelligence), and a further distinct property named by 'intelligence'. It is this latter name which is to be defined as synonymous with a certain description, and which picks out, as it happens, IFS. But the name for which property is it that he is defining as synonymous with a certain description, if not the property we attribute when we say 'John is intelligent'? There is no other property around that I can see. I am not here questioning--it should be clear--Lewis' right to define terms the way he does; I am pointing out that in so defining, e.g., 'intelligence', it does not make sense to say that what is being defined is a term for something other than what is at-

tributed in 'John is intelligent'. Recall that it is our everyday meanings that Lewis is after. It would not, therefore, do to just establish by fiat that 'intelligence' is to be construed with a new meaning, not naming that which is attributed. Of course, Lewis could say this, but then he would have to give up, it seems to me, his claim to be analyzing our everyday meanings; for in that context, intelligence is what we attribute in 'John is intelligent'.[5]

As mentioned, Lewis reformulates theories so that all theoretical predicates are eliminated in favor of names. His motivation for this is convenience and esthetics: it eliminates the necessity of having more than one type of variable, as in Ramsey's treatment (as well as Carnap's). The implication (and perhaps explicit assertion) is that it should make no difference whether we have names or predicates; and this seems as it should be, for one imagines that it should not make a difference which of two logically equivalent formulations one uses.

So let us leave the predicates in; let our theory be, 'If John is in pain, then John is apt to shout', with our theoretical (psychological) predicate, ' . . . is in pain'. The most natural way to form a defination à la Lewis would be ' $\exists F($ F(John), then John is apt to shout $)$'. But the

---

[5]Actually, an analogous point can be made if we consider standard theoretical predicates, and not just those of folk psychology.

only property we have here defined is the property of being in pain--that is, the property attributed to John. And we know that this is not the property Lewis wishes to define.

But can we not isolate 'pain' and define it? Perhaps our definition should be, '$\hat{F}(x)$ (John is in x and Fx, then John is apt to shout)'; but we have now defined the property of being a pain, a property attributed to states. This is equally of no help, for ' . . . is a pain' is not synonymous with (say) ' . . . is a CFS'; we could not allow these predicates to express the same property.

Indeed, whatever term we now define, will be, by assumption, in predicative position. And we will thus always have nonsynonymous predicates attributing the same property (contrary to Lewis' wish): the defined predicate, whatever it is, and the predicate for the corresponding physical property. So it does make a difference, it appears, whether we use names or predicates; Lewis seems to be able to hold all his views only if he sticks to names.

Finally, let us consider the postulate of folk psychology with the psychological terms replaced by variables. This postulate must, according to Lewis, be uniquely satisfied (or nearly so) if our folk psychology is true. Letting 't' be the variable which replaces 'pain' in our theory, consider the interpretation which purportedly uniquely satisfies the postulate, but with the exception that we assign to 't' not CFS but the property of having

pain, which property, let us recall, is coextensive with pain in the actual world. This new interpretation preserves truth; the only part of the postulate which is affected is, let us say, '(x)(If x has t, then x is apt to shout)'--which will remain true. So now there are <u>two</u> interpretations satisfying the postulate, contra hypothesis.

Let me conclude by saying once more that I am cognizant that this last problem, along with the previous ones to varying degrees, is an artifact of Lewis' formulation and is possibly resolvable by changing the formulation (in this case, perhaps by insisting that one needs unique <u>physical</u> realizations; this is in Lewis' spirit, but not strictly speaking, in his letter). They remain, nevertheless, problems Lewis must deal with.

In the remainder of this section, I shall suggest and consider a way out of the above problems, including that of the nonphysicality of having pain. The basic idea will be to distinguish the senses of predicates from what they refer to--i.e., essentially a Fregean conception. This will allow nonsynonymous, nonnecessarily coextensive, predicates to refer to the same property. (I do not mean to prejudge by my use of the term 'refer' here, whether or not the relation that holds between singular terms and objects is the same as that which holds between predicates and properties; I just know of no better term.)

There are various ways one might attempt to formulate this. We might, for instance, distinguish universals,

the reference of predicates, from properties, their senses, requiring only actual (and not necessary) coextensivity from universals. The fact then that ' . . . is in pain', and ' . . . is in CFS' are not synonymous would not prevent them from referring to the same universal, and one would not have to make being in pain nonphysical. Of course, on this conception, we must have some identity criteria for universals: perhaps nomological equivalence would do.

The method I prefer for accomplishing our Fregean objective is perhaps best presented by introducing the notion of nonrigid predicator.[6] An example of this would be ' . . . has the property t' where 'the property t' is a nonrigid designator--e.g., 'the property I am not thinking of'. If 't' refers to CFS in this would and DFS in world w, then ' . . . has the property t' will refer to having CFS in this world and having DFS in world w.[7]

---

[6]This way fits in more smoothly, I believe, with Lewis' actual views: we need not distinguish, as Lewis does not, two sorts of nonparticulars, universals, and properties. Further, the criterion of nomological equivalence would not fit in easily with Lewis, for we could not then identify pain with any physical universals, since pain, as we have seen, is not even coextensive with any one such universal.

[7]Montague (in, "On the Nature of Certain Philosophical Entities," in Formal Philosophy) has an argument which seems to show that where t is a nonrigid designator, its referent is not, generally, identical with the referent of 'the property of having t'. If examined closely, it will be seen that his argument depends on 'The property of having t' being a rigid designator. If it is not, then his argument does not follow.

He does further show that we can not deduce that the experience Jones has (his example) is identical with the property of having the experience Jones has, from logically true premises. But he does not show that we can deduce that they are not identical either! And if 'The property of having the experience Jones has' is nonrigid, then they can be identical.

There is nothing in the above that I can see which would prevent us from taking the properties having CFS, and having DFS, from being interpreted as standard Lewis set theoretic constructs. And if we wish to stay as close as possible to Lewis, we can say the same for the senses of predicates. Thus, in terms of Lewis' formalism, my suggestion would be that where t is nonrigid, the sense of 'The property of having the property t' (or, ' . . . has the property t') is the function which assigns to any possible world w the property Q_w, where Q_w has as its extension in any possible world z those things which have, in z, what t names in w. This function will be different from what is referred to by the predicate (e.g., see immediately below).

The above formulation gains the physicality of having pain. For ' . . . having pain' just like 'pain' will refer to CFS: it refers, by the above, to having CFS, but that, according to Lewis, is itself identical to CFS (assuming, as we shall, that 'CFS' is rigid). And, once again, this fact need not be reason to worry over the nonsynonymy (and possible noncoextensivity) of the (senses of) predicates ' . . . is in pain', and ' . . . is in CFS'.

It should not need pointing out that the above suggestions are incompatible with Lewis' views as they now stand. For him, properties (attributes, states, etc.) are senses of predicates, and determined by necessary coextensivity. The view I am suggesting is, however, compatible with the spirit

of Lewis' doctrines (the physicalist ones anyway); and it would do away with the problems presented earlier. Further, one would think it a natural move on Lewis' part: He allows that the names 'pain' and 'CFS' can refer to tho same state though they have different senses; why not similarly for the predicates, ' . . . is in pain' and ' . . . is in CFS'?

There are also independent reasons for thinking that something like the above is correct; I shall cursorily present some of them.

The first point has really just been mentioned, but let me repeat. We allow names with many different senses to pick out a single referent, why not similarly with predicates? If we ask after the truth of a given sentence, we, typically, think of an object which may have been variously referred to, and determine if it, in fact, has the property attributed--and it seems natural to add, a property which may have been variously attributed (referred to).

For the next point, imagine that I am now thinking of the color blue. In these circumstances, it seems true to say that the table's being the color I am thinking of just is this table's being blue: there is here one event variously described. In a different context, this table's being the color I am thinking of might have been this table's being green.[8]

---

[8]In a similar vein, Kim has said, "There is a sense in which his being a good man consists in, amounts to, and is nothing over and beyond, his being generous, sympathetic, and truthful, . . . it is, rather, that being a good man represents, or consists in, different properties in different instances." Kim, "Causality, Identity, and Supervenience

One possible analysis of these facts--though by no means the only one!--is that in the present circumstances, ' . . . is the property I am thinking of' and ' . . . is blue' refer to the same property.[9] We may, to be sure, have used different descriptions with different senses to refer to this property; and, consequently, in different circumstances, ' . . . is the property I am thinking of' might refer to a different property than ' . . . is blue'; but in the circumstances that actually obtain, they do refer to the same property; and this is just to say that ' . . . is the property I am thinking of' is a nonrigid predicator.

To the extent that my above analysis is persuasive, and I admit that there are other reasonable competitors, it supports the view I am defending.

The next point has already been briefly mentioned. It seems natural to say, following Putnam, that, "the physical property of having a particular temperature is really (in some sense of 'really') the same property as the

---

(Fn. 8 continued)
in the Mind-Body Problem," in Midwest Studies in Philosophy Volume IV, Studies in Metaphysics. French Uehling, Wettstein, eds. This point is not uncontroversial.

[9] I should point out that a more standard analysis (e.g., Kim's) would have it that in the present context, the particular instances of the two properties, though not the properties themselves, are identical. I do not know whether this analysis can be made compatible with the view I am concerned to defend.

property of having a certain molecular motion."[10] This seems prima facie correct, even though the corresponding predicates are not synonymous (or necessarily coextensive).

Actually, the view Putnam expouses is very close to the first one we mentioned above. On Putnam's view, predicates (not linguistic entities) correspond to what I called there "properties"; and physical properties correspond to what I called there "universals." The difference in the views lies in the difference between the identity criteria for physical properties. For Putnam, A = B if A reduces to B, or B reduces to A, or both reduce to C (Putnam's article is concerned with developing this point). In any event, the important point, as far as we are concerned, is that he exhibits a credible example of nonsynonymous predicates referring to the same property.

Finally, there seems to me to be more mundane (than the above) cases of nonsynonymous predicates referring to the same property.

Perhaps the best way to approach what I have in mind is by way of extension of Donellan's referential/attributive distinction to predicates. I shall not go into details, since they are widely known, but briefly, the standard distinction is this: If at a party, someone remarks, "The man in the corner with a martini in his hand is a

---

[10] Putnam, "On Properties," Mathematics, Matter, and Method.

genius," he may have succeeded in referring to a particular individual, though in fact, there is no man in the corner with a martini (his glass contains water). This is a referential use of the referring term 'The man in the corner'. An attributive use is one in which we take the description to refer to whatever, if anything, literally satisfies the description; in the above example, no one.

Let us now extend this distinction to predicates. I say to you, "John has the most dangerous character trait possible for a politician." I might just mean by this that John has whatever character trait it is which is most dangerous--about which I need have no opinion: this would be an attributive use of the predicate. On the other hand, we may be just concluding a long political discussion in which I strenuously contend that lust for power is the most dangerous character trait possible for a politician; by way of example I say, "John has the most dangerous character trait possible for a politician." This is a referential use of the predicate, for I have clearly attributed lust for power to John, <u>whether or not</u> lust for power is, in fact, the most dangerous character trait possible for a politician. (I certainly have not attributed stupidity to him if, say, that happens to be the most dangerous trait possible for a politician--I may think he is brilliant.)

This relevance of all this to our present concerns is that it is an everyday case of nonsynonymous predicates

referring to the same property: ' . . . lusts for power', and (used referentially) ' . . . has the most dangerous character trait possible for a politician'.

Another approach to the above point is to invoke Lewis' first reading of 'the property of having t' where t is nonrigid (see quote on pg. 33). On the second (and proprietary) reading, the property attributed by, e.g., ' . . . has the most dangerous character trait possible for a politician', is that which has as its extension in any possible world w, whatever has the most dangerous character trait possible for a politician on w: a property which is distinct from lusts for power.

But it seems to me that in the situation described above, the first reading would be more appropriate. On that reading, the property is that which has as its extension in any possible world w, the extension of the most dangerous character trait possible for a politician in the actual world--i.e., lust for power (say). For lust for power is, it seems to me, the property being attributed. And again, if this is correct, we have nonsynonymous predicates attributing the same property: ' . . . lusts for power', and (on Lewis' first reading) ' . . . has the most dangerous character trait possible for a politician'.

This latter approach has problems the attributive/referential approach does not.[11] Thus, the example I have

---

[11]The referential/attributive distinction for predicates is not the same as Lewis' two readings of 'the property of

used might strike the reader as also one in which Lewis' second reading is more appropriate. Further, choosing between Lewis' two readings seems, to a large extent, to be a matter of convention. Even so, I would assume that the reader, understanding what I am driving at, will be able to construct what, by his lights, is a more suitable example: one in which it is not just a matter of convention what the right reading is. (After all, it isn't <u>just</u> a matter of convention what sense predicates have in a given situation).

Another problem is that Lewis might say that even if he were to grant that, in this case, the first reading is more suitable, he could still claim that <u>in those circumstances</u>, the predicates, ' . . . lusts for power' and ' . . . has the most dangerous character trait possible for a politician', are not only coreferential but also used synonymously.

If anyone were to adhere to such a view, there would be no way that I could think of to refute him. But the view does seem prima facie false; even in the above contexts, the predicates clearly do not have the same sense. The only possible reason for claiming that they do would be that it

---

(Fn. 11 continued)

having t'; just as the normal referential/attributive distinction is different from the rigid/nonrigid distinction. Indeed, there is a very close parallel between singular terms and predicates in this respect. Thus, both of Lewis' readings can be construed either referentially or attributively.

is a consequence of Lewis' other views; but since it is just the reasonableness of those views which are presently in question, appeal to them can hardly be countenanced.

Whether we use the referential/attributive distinction, or the more problematic one of Lewis' two readings, the main point remains the same: we have an example of non-synonymous (i.e., whose senses are not necessarily coextensive) predicates referring to the same property.

These considerations can, of course, hardly do justice to the complex issues they deal with--and they were not meant to. Their purpose is just to lend some initial credence to the suggestion that we distinguish properties from the senses of predicates.

And the point of <u>that</u>, as far as we are presently concerned, is to suggest a modification to Lewis' views which would allow him to deal with the problems presented previously; most importantly, it would allow him to deal with the nonsynonymy of predicates without making, e.g., having pain, nonphysical (as he presently does). Doing this would, it seems to me, make Lewis even more of a physicalist.

Let me conclude by emphasizing once more that though the above modification may be appealing, it is not one Lewis makes or suggests. As his views presently stand, the problems mentioned remain to be dealt with.

SECTION IV

IV

We have, until now, assumed the correctness of Lewis' analysis of theoretical definition and reduction. Thus, in the first section, we took it as a virtue of Lewis' views on everyday psychology that they were subsumable to the more general theoretical case. In this last section, we shall turn our sights on this more general case. Since the considerations adduced will usually be applicable to the case of psychology as a particular instance, they will have, besides their intrinsic interest, implications for our previous discussions. Indeed, the issues here raised will often parallel already mentioned ones.

The structure of this section will be somewhat similar to that of section III. The first part will concern itself with one general problem of considerable significance; the second, with a few problems of a more technical nature.

My first concern is with the dogma--used as a premise in Lewis' reductive arguments--of, "The explanatory adequacy of physics," or "The Unity of Science." (Recall the discussion in section I.) I.e., that (à la Putnam and Oppenheim) all true theories are ultimately reducible to physics. Thus:

> My second premise is the plausible hypothesis that there is some unified body of scientific theories, of the sort we now accept, which together provide a true and exhaustive account of all physical phenomena (i.e., all phenomena describable in physical terms). They are unified in that they are cumulative: the theory governing any physical phenomenon is explained by theories governing phenomena our of which that phenomenon is composed and by the way it is composed out of them. The same is true of the latter phenomenon and so on down to fundamental particles or fields governed by a few simple laws, more or less as conceived in present-day theoretical physics. I rely on Oppenheim and Putnam for a detailed exposition of the hypothesis that we may hope to find such a unified physicalistic body of scientific theory and for a presentation of evidence that the hypothesis is credible.[1]

Further, all theoretical terms, he contends, will refer, when functionally defined as he suggests, to physical entities--given the above dogma. He does allow, as we know, for the possibility of nonphysical entities; but he does not allow that they are causally efficacious with respect to any phenomena within the domain of any scientific theory.

It is this I wish to consider; again, I will be concerned only with nonparticulars (properties, states, experiences, etc.). As a test case, a particular example will be used: that of genes and DNA molecules. I choose this example because it is a <u>particularly good</u> case for Lewis. For it is as good an example as there is of a theoretical term which is, on the face of it, definable by its causal role.[2]

---

[1]"An Argument for the Identity Theory," Pg. 169.

[2]It has also been claimed (by Armstrong and others) to be a paradigm of what psychological reduction will look like.

If the reader feels that there is something peculiar about my example which makes it inappropriate, he should substitute his own in the following arguments (e.g., 'electron' or 'quark').[3]

Further, it would seem, at first blush, that this case fits neatly into the pattern of Lewis' analysis. First, it is true that genes are DNA molecules; and secondly, the apparent difference between the universals (properties, whatever) might plausibly be attributed to the difference between the concept 'Gene' and the concept 'DNA molecule'. We could thus account for our intuition that the two might not have been identical.

In discussing this example, I shall first restrict myself to Lewis' actual views; I shall then turn to Lewis' views as modified in the way suggested in the previous section. The problem I am concerned with can be brought to the fore by asking, "What universals, in the present case, is Lewis identifying?"[4]

---

[3]Of course, Lewis' views are meant to apply to all theoretical terms; so even one counterexample would be sufficient. But a stronger claim is intended. I am taking this as a particularly good example for Lewis. If there are problems here, there will be problems generally.

[4]Carnap (See: An Introduction to the Philosophy of Science) translates everything into a language which consists of just theoretical class and relation terms--e.g., 'Mol' is to stand for the class of molecules. He then proceeds in the usual manner to define the theory's Ramsey sentence. Whatever the merits of this strategy, it is of no use to Lewis. The most it would give us is that the class of pains (say) is identical with the class of CFSs--i.e., token/token identity.

In the case of everyday psychology, pain was identified with CFS. This was accomplished, essentially, by positing the trichotomy:

1. Token pains
2. Pain
3. Having pain

This is not unreasonable in the case of experiences such as pain.[5] When, however, we come to the present case, and in the realm of objects talked about by theories generally, we seem to be more restricted; there is only:

1. Genes (electrons)
2. Being a gene (being an electron)

There seems to be nothing in these cases corresponding to pain.[6] So, whereas in the pain case Lewis had two terms, one of which he could interpret rigidly and the other ('pain') nonrigidly, identifying the latter with CFS, an analogous move does not seem possible here; there seems to be no clear way of extending Lewis' views to theoretical terms generally.[7]

---

[5]Though we should keep in mind that even in these distinctions, Lewis is atypical. For him, the extension of pain consists of individuals, where most would take it to consist of states. (Do not confuse this with the fact that for Lewis, the reference of 'pain' is a state.) Thus, we could really add for Lewis a fourth to the above list: the property of being a pain.

[6]Or if you identify pain with being a pain (See Fn. 5), there will be nothing corresponding to having pain.

[7]The formulation of the point just presented is due to Ned Block.

How can Lewis respond to this objection? There is only one way I can think of. (Since Lewis does not directly deal with this question, one can't be absolutely positive, but (1) I can not think of any other response he might make; (2) The response I give parallels his discussion of pain; and (3) Though not directly stated, the view below is strongly implicated by his discussion of theoretical terms.)

He would respond by positing in, e.g., the present case, distinct from both the set of genes and the property of being a gene, a nonparticular referred to by 'Gene'; this term, he would say, is defined as synonymous with a certain nonrigid functional description. It would probably be most appropriate to refer to such a nonparticular as a "kind."[8] If we follow the analogy with the pain case all the way, we would now have 'the property of being a gene' as a rigid designator which names in every possible world the property which has as its extension in any world w_1, the extension of what 'gene' names in w_1.

This response is vulnerable to a charge of implausibility--a charge of considerable force, I believe. There just does not seem to be some further universal other than the property of being a gene which might be referred to by 'Gene'. (And as we have seen, there are those such as Davidson who not implausibly contend that terms like 'gene' are not names for anything.)

---

[8]I do not know whether Lewis would use "kind" terminology; the precise terminology does not, of course, matter; the point would remain the same.

On Lewis' behalf, we may respond that taken from his particular perspective, there is nothing particularly problematic about kinds as construed above. They are on par with universals (properties, attributes, experiences, etc.) in being functions from possible worlds to extensions. In the gene case, as in the pain case, he is merely distinguishing two such functions. One, in the present case, named by Θ and the other named by 'the property of being Θ'. These two could be identical only if Θ is rigid; if it is not, and theoretical terms standardly are not according to Lewis, the two will not be identical. The kind I am thinking of may be Horse, but this is not what I attribute when I say, "Affirmed is an object of the kind I am thinking of."

A response such as this on Lewis' behalf will, obviously, be open to objections analogous to those to the last section. Instead, however, of reiterating, I shall focus on what I take to be the crucial problem with such a view. Let me say that while I am confident that the view I am attacking is Lewis', for the discussion below it is not crucial that I have gotten every detail correct. What is crucial, it appears, is Lewis' explicit contention that theoretical terms be nonrigid.

The response we are attributing to Lewis might be thought to have some independent support. Thus, we often hear it said that, as Fodor puts it, "The natural kind

predicates of a science are the ones whose terms are bound variables in its proper laws."[9] That is, if we have a proper law of the form: $(x)(Px \rightarrow Fx)$, then the class of Ps form a natural kind. The variables in such a law are, in a sense, taken to range over possible as well as actual objects; for we intend the law to assert of any possible object that if it were a P, then it would be an F: We mean to assert more than just that all actual Ps are Fs. The natural kind P is then plausibly construed as determined by saying which things have P in any possible world.

Consider for a moment 'Horse' and other such natural kind terms. Typically, they are count nouns, sortals, or whatever (e.g., mass terms). In most contexts such as 'There is a horse over there', they can not be said to refer to anything at all; or if they do refer, not as singular terms do.

If we do construe such terms as names, then there is a natural candidate for what they refer to: the kind, in the case of 'Horse', of all those things actual or possible (or the appropriate function) which are horses--that is, which have the property of being a horse. If we are inclined to identify kinds with properties, this would just be the property of being a horse. Terms such as these let us call 'natural kind predicates', and say that they express when

---

[9] Jerry Fodor, "Special Sciences," Synthese 28 (1974): 97-115.

used predicatively, the kind we say they name.

We can now rephrase the above Fodor quote as the contention that the natural kinds of a science are those which are expressed (in the above sense) by the natural kind predicates of its proper laws. If P is such a predicate, then the class of Ps, actual or possible, form a natural kind of that science. And the laws of that science purport to make true generalizations over just those kinds.

We can now see what is wrong with Lewis' view. On it, the theoretical terms for kinds or properties which occur in his formulation of theoretical postulates, do not name the kinds or properties which, in the above sense, the theory purports to make true generalizations about; for the theoretical terms are not natural kind terms in the above sense: if, according to Lewis, Θ is a theoretical term naming a kind, it will not, since it is nonrigid, generally name the kind which has as its extension in any possible world those things which have the property of being a Θ. Thus, 'Gene' would name the kind DNA molecule whose extension in any possible worlds does not consist of those things which have the property of being a gene.

To be sure, Lewis does countenance terms for kinds of properties which satisfy my above conditions, to wit 'The property of being a gene'. But this is not the theoretical name which gets functionally defined; it is defined, according to Lewis, but in such a way that it is <u>not identical to</u>

any physical property or kind.[10]

Thus, it is that Lewis has not defined so they come out physical the kinds over which laws purport to make true interesting generalizations. The laws of (statistical) genetics purport to make true assertions which support counterfactuals about anything that might be a gene--not just DNA molecules (which is in the domain of genetics when construed as a part of biochemistry). And it is <u>these</u> kinds, it would seem, which it is incumbent upon a physicalist to show physical.

While from one perspective this objection is indeed strong, it can, again, not be pushed too far. All I have really shown is that Lewis' views are inconsistent with others; I have not shown him to be internally inconsistent. Different is not necessarily wrong; and if Lewis' analysis is incompatible with the standard one, no conclusion can be drawn on this basis alone.

Nevertheless, we can, I believe, say that Lewis' analysis does not capture the motivation behind physicalist (type/type) doctrines. A large part of that motivation, as I understand it, is to show that all natural kind terms of the special sciences reduce to physical kinds. And this, I believe, he has not done.

---

[10]Thus, it is not identical with the property of being a DNA molecule, nor with being a RNA molecule. It will have a different extension in some possible world from every physical property or kind. (Compare with the property of having pain.)

The mere fact that all theoretical terms on Lewis' formulation refer to physical entities should not be given too much weight. We could, as already mentioned, reduce all properties by identifying them with the set theoretic functions to their extensions; and then claim set theory as part of the physical sciences. Would this be a satisfactory reduction? Or we might eliminate all theoretical terms from our laws by Craig's strategem, and hence, have nothing left to reduce. Would this be a satisfactory reduction? No and no. Similarly, the mere fact that all theoretical terms in the formulation refer to physical entities (which is true for Craig also) does not, in itself, make for a sufficient reduction. More is required, and I think we will not find it in Lewis.[11]

It is of some interest to see how our test case would fare if Lewis' views were modified as suggested in the last section. This is, by introducing a property/universal (predicate/physical property for Putnam) distinction; or by introducing my notion of nonrigid predicator.

It would seem that in this case, my query, "What universals should we identify?", will have a very simple response. We can identify the property of being a gene with the property of being a DNA molecule, and not worry that the corresponding predicates are not synonymous.

---

[11] I should lay my cards on the table and say that I believe no formulation can do it.

The simplicity of this response is deceptive; consideration will show that these two properties are really distinct. Being a gene is being the satisfier of a certain causal role, and that is not the same property as being a DNA model. This is not due merely to the nonsynonymy of the two predicates (if that were the only reason, we would be begging the question); it is, rather, because we can think of nomologically possible worlds (worlds in which the laws of nature are as they actually are), perhaps even the actual world, in which the entities having the property of being the satisfier of the appropriate causal role are not DNA molecules; they might be RNA molecules or miniature ropes and pulleys, etc. Given this, the property of being a gene could not be identical to the property of being a DNA molecule.

This point is similar to the standard antireductionist counter to type/type physicalism. As such, it is, as are the standard counters, too strong as it stands. I certainly have not proven my point. For, to claim nonidentity on the basis of possible nomological noncoextensivity is, in certain respects, to put the cart before the horse. What is true is that, if we were to identify being a gene with being a DNA molecule, then the properties would be coextensive in all possible worlds; not that we can, because properties seem noncoextensive, infer nonidentity.

To see that this is the correct way of looking at

things, consider that it would clearly be a bad argument, since the conclusion is false, to contend that the property of having a certain temperature could not be the property of having a certain kinetic energy; and similarly with the property of being a body of water, because they are not coextensive in all possible worlds; for since water (say) is H_2O, any world in which water occurs is a world in which H_2O occurs.

From our present vantage point, with our knowledge of the identity of water and H_2O, we can make the above reasonable response. What this shows is that prior to our knowledge of these identities, it would have been invalid to argue that being water and having a certain temperature could not, because of possible nomological noncoextensivity, be identical to being H_2O and having a certain kinetic energy. Similarly, we can not now eliminate the possibility that the property of being a gene is identical with the property of being a DNA molecule merely on the grounds of the supposed possible nomological noncoextensivity of the two. Whether we construe the two as possibly noncoextensive or not depends on whether we construe them as diverse or not; and not, as is often implied, the other way around (as if one could just look in other possible worlds and see).

So I have not shown that Lewis, as modified, would be wrong in making the identification. Nevertheless, it remains pretty clear that they are not: for the gene/DNA molecule

example is not completely analogous with the temperature/ kinetic energy case. We already pretty well understand the relation between genetics and the biochemical entities and mechanisms which underlie it, just as we understand the relation between statistical mechanics and thermodynamics. It is safe to assume that (barring a complete paradigm change) no new discoveries will have any effect on our understanding of these relations. And <u>still</u> we are inclined to say that there could <u>not</u> be temperature which was not kinetic energy, while there <u>could</u>, even nomologically, be entities that had the property of being genes which were not DNA molecules; there could be other things which functioned like a gene, and to function like a gene is to be a gene (though to feel, look, etc., like water, is not to be water).

Essentially, this same point can be made using Kripkean terminology. It is metaphysically necessary that: temperature= kinetic energy; while it is not metaphysically necessary, even if we restrict ourselves to nomologically possible worlds, that: DNA molecules are identical to the satisfier of the appropriate causal role. The contention that the first of the above claims is false because we can imagine cases in which temperature turns out not to be associated with kinetic energy, can be dealt with by appeal to Kripke's notion of epistemic necessity.[12]

---

[12]This view is discussed in detail in the portion of the thesis which deals with Kripke.

So in the end, Lewis' views, as modified, do fall prey to the standard antireductionist argument; though we have seen that that argument is not as strong as is often supposed. The real Lewis sees this problem, in a sense, and deals with it by eliminating reference to the recalcitrant properties from theoretical postulates. Victory for him is obtained by having all theoretical terms in his formulation refer to physical entities; he is not concerned about any other entities which escape his net.

But this is, from a certain perspective, a Pyrrhic victory indeed. Eliminating reference to entities does not eliminate those entities. Lewis needs, by my lights, to explain, or explain away, the appearance that he does not reduce the universals which, as explained above, the postulates of special sciences are committed to. Or if he thinks such postulates are not, in fact, committed to those universals, explain why he thinks that.

I turn now to various, somewhat more technical problems. While these do involve the particulars of Lewis' formulation, they seem to me to be, moreso than, say, the problems of section three--problems of principle more than problems of detail.

The form of theoretical definitions for Lewis is essentially that which was given for psychological terms at the beginning of section one; I shall not bother repeating it here. His treatment of theoretical reduction is likewise

essentially that an informal instance of which was given at the beginning of the discussion of Lewis. It will, however, pay for us to have a more precise formulation:

> It may happen after the introduction of the T-terms, that we come to believe of a certain n-tuple of entities, specified otherwise than as the entities that realize T, that they do realize T. That is, we may come to accept a sentence
>
> $$T(r)$$
>
> where $r_1 \ldots r_n$ are either O-terms or theoretical terms of some other theory, introduced into our language independently of $t_1 \ldots t_n$. This sentence which we may call a weak reduction premise for T, is free of T-terms. Our acceptance of it might have nothing to do with our previous acceptance of T. We might accept it as part of some new theory; or we might believe it as part of our miscellaneous, unsystematized general knowledge. Yet, having accepted it, for whatever reason, we are logically compelled to make theoretical identifications. The reduction premise, together with the functional definition of the T-terms and the postulate of T, logically implies the identity:
>
> $$t=r$$
>
> In other words, the postulate and the weak reduction premise definitionally imply the identity $t_i=r_i$. Or we might somehow come to believe of a certain n-tuple of entities that they uniquely realize T; that is, to accept a sentence
>
> $$\forall x(T(x) \equiv x=r)$$
>
> where $r_1 \ldots r_n$ are as above. We may call this a strong reduction premise for T, since it definitionally implies the theoretical identifications by itself, without the aid of the postulate of T. The strong reduction premise logically implies the identity
>
> $$r=\iota xT(x)$$
>
> which together with the functional definition of the T-terms, implies the identities of $t_i=r_i$ by transitivity of identity.[13]

First point. There has been considerable discussion on what constitutes lawlikeness, or lack thereof, in a given

---

[13]"Psychological and Theoretical Identifications," Op. Cit., Pg. 255.

theoretical statement. About the only criterion about which there seems to be general agreement is that a statement, whether true or not, is lawlike if it supports counterfactuals. Lewis' formulation, if what I say below is correct, vitiates this criterion. There are many statements which, if taken as part of a theory, do support counterfactuals on that formulation, though they are clearly not lawlike, and would not, on a standard reading, be said to support counterfactuals.

Consider, 'All coins in my pocket are copper'. Suppose I add this statement to the laws of my new chemical theory. A standard, and legitimate, criticism of my theory would be that, whether true or not, it is certainly not lawlike, for it does not support counterfactuals: if there were dimes in my pocket, they would not, in point-of-fact, be copper.

But suppose I adhere to Lewis' views and am construing 'copper' as a theoretical term. Then it will be defined, forgetting for the moment about the rest of the theory, as '$\eta x(y)$(y is a coin in my pocket $\rightarrow$ y is x)'. What now of a (nomologically possible) situation in which there are dimes in my pocket? If we accept Lewis' account then, given the above definition, these dimes would be copper in that situation; for 'copper' is a nonrigid designator, referring in any possible situation to whatever fills the appropriate

conditions--in this case, some dimes.[14] So the comment that "'All the coins in my pocket are copper', is not lawlike because if the coins were dimes, they would all the same not be copper," can not be made by Lewis; they would be copper--as defined by such a crazy theory--on Lewis' view. Now, granted the theory being considered is silly and not one ever seriously proposed. My point just is that according to Lewis, there is no way to sift out these crazy theories (on the basis of counterfactuality) from the reasonable ones.

What underlies this problem is, I think, the semi-analytic nature of Lewis' theoretical definitions: in any possible world in which the terms refer the statement is true. Thus, as Hemple in another context has said:

> For if the principles asserted by a scientific theory are implicit definitions of its key terms and hence analytic, the role of experiment and the need for empirical evidence are thrown into question. If--to construct a schematic example--Galileo's Law and Kepler's Law are taken to be definitive of 'free fall' and of 'planetary motion', then there would be no need for experimental observations test. Moreover, empirical data on the actual motion of the planets about the sun would be irrelevant to those laws. If the findings did not conform to the law, this would show only that actual fall is not free fall as implicitly defined by Galileo's formula or that the actual motion of the planets is not planetary motion as implicitly defined by Kepler's laws. The laws would be analytic; in order to make them applicable to their usual empirical subject matter and thus to restore the relevance of empirical testing, they would have to be supplemented by laws to this effect: the fall of a body in a vacuum near the surface of the earth is free fall

---

[14]This counterfactual will come out true on Lewis' analysis; see his, Counterfactuals (1973).

> as characterized by Kepler's laws. But the 'theories' obtained by such supplementation clearly are no longer analytic; their terms are no longer implicitly defined by them.[15]

As it stands, this quote is not directly relevant to Lewis. For Lewis, O-terms just mean old, already understood terms. Consequently, these terms could very well suffice to tie down theoretical postulates and definitions to particular interpretations. One of the O-terms might, for instance, be 'near the surface of the earth'.

But we can modify the above quote and put in the form of a dilemma so that it does present a problem for Lewis. We can say that either (1) the O-terms are so general as not to tie down the theory to any particular interpretation, in which case the above Hemple remarks apply; or (2) the O-terms do tie down the theory to a particular interpretation; in this case, however, the definition of the theoretical terms will, it seems, become implausible. Do we really want 'near the surface of the earth' as part of the theoretical postulate which consists of Kepler's or Galileo's Laws? I think not; but it would have to be there, it seems, if this horn or the dilemna were grabbed.

The strength of such an objection is diminished by (1) the possibility that Lewis would contend that such terms as 'near the surface of the earth' are (at least implicitly) in the theories; and (2) there being no eliminating the

---

[15] Hemple, Carl. "Reduction: Ontological and Linguistic Facets." In, S. Morgenbesser, P. Suppes, and M. White (eds.) Essays in Honor of Ernst Nagle, Pp. 191-192.

possibility that Lewis' formulations could find a happy medium: with O-terms specific enough to tie down the interpretation, but not so specific as t make the definitions implausable. Perhaps this is so, but I think if Lewis wishes to defend one of these positions, the burden of proof will now fall upon him to show them reasonable.[16]

---

[16]Here is another objection, which I offer only tentatively:

Consider an unreduced theoretical term such as 'quark'. Given a standard nonLewis functional account of theoretical terms, we might propose the following problem. Consider a possible situation in which little men satisfy the appropriate functional definition (For the details of such an example, see Block, "Troubles with Functionalism"). The existence of such an example would seem like good reason to deny the synonymy of, in this case, 'quark' with the appropriate functional definition (and similarly, with 'pain'); there are, after all, possible situations in which the description picks out little men and not quarks (or pains).

We know that this problem will not be applicable to Lewis' views: for on his account, 'quark' is nonrigid and can pick out different things in different possible worlds (such as little men). The ability to make this response is a substantial virtue of Lewis' treatment, one that Lewis is at pains to emphasize; but it is a knife that cuts both ways.

For if 'quark' does not purport to pick out one thing in every possible world, what can it be said to refer to in the actual world? How do we fill in the following blank: In world w_1, 'quark' picks out little men, but in the actual world, it picks out_____? The natural response seems to be quarks (or the synonymous description), but that will not do; for quarks may be very many things according to Lewis, including little men; what is it in the actual world?

The temptation is, no doubt, strong to insist that all one need say is that quarks are whatever satisfy such and such a causal role and leave it at that. But recall that we are considering unreduced terms; we have still not said anything which will show us that quarks are not actually little men.

The only way out of this, as far as I can see, is to introduce some ostensive component beyond the description: getting in front of an instrument and saying, 'What caused that is a quark'. But of course 'quark' will now now be, contra Lewis, analytically defined by the original description.

(Fn. 16 continued)

Again, as in the problems mentioned above, it is the analyticity of Lewis' definitions which seems to be the culprit.

There are a few possible responses which might be made to all of this. First, it might be claimed that Lewis' functional definitions are not meant to apply to unreducible theoretical terms. I do not think Lewis would, or reasonably could, respond thus. His writing suggests that he means his analysis to apply to all theoretical terms. Further, his definitions are supposed to be analytic. But where reduction stops, which theoretical terms, if any, are ultimately primitive, is an empirical question; and if the definition of a term hinges on an empirical question, it can not be analytic. And in any case, such a restriction on Lewis' part on what gets functionally defined, if he were to make it, would be ad hoc. Lewis' definition is of the sort which one expects to hold generally if it holds at all. Indeed, the most plausible examples of theoretical terms functionally defined often are those which are unreducible (e.g., our 'quark' case); it is just in such cases that we commonly reify entities and give them names in order to fill some causal role.

It might also be contended on Lewis' behalf that for him, theoretical terms do not really name anything--they are theoretical terms à la instrumentalism. Lewis explicitly denies that this is his intention.

Finally, it might be said in Lewis' defense that he could just invent a name 'Krauq', defining it as referring (in every possible world) to what 'quark' refers to in the actual world. My question could now be answered by saying that 'quark' picks out Krauqs in the actual world. This will also not do; the reference of 'quark' is still untied. (We have violated what Kripke calls the noncircularity condition on reference fixing.) For we are here saying no more than 'Krauq' refers to whatever 'quark' actually refers to; given that there is no independent way of saying what that now is, we are no further out of the woods.

Of course, we could always say kraqus are <u>that</u>, pointing to some instrument reading. But this will just bring back the ostensive component into the definition, and be of no additional help.

Though I am not entirely convinced of the validity of the objection I am presently raising against Lewis, if it is valid, then it is very strong. It shows that--at least for a large subset of theoretical terms--we must have an ostensive component.

For my next point, let us consider a theory, T, which is reduced to a more basic theory, R. Let 't' be some theoretical term (for a universal) and T (defined à la Lewis) and 'r' the term of R which, on the basis of the reduction 't', is reduced to; so that 't' and 'r' are coreferential. Instead of T, we also write T(t).

Part of Lewis' scenario for reduction is that we discover that T(r); that r, something referred to by a theoretical term of the reducing theory, satisfies the causal role definitive of 't'. And it is very often, though not always, the case that T(r) <u>follows</u> from the reducing theory r. Such cases, contends Lewis, are the significant cases of reduction. So that this point is clear, I will quote at some length:

> Suppose that during this period, T is reduced by means of some other accepted scientific theory T* . . . . The more interesting case, however, is that in which T* is well systematized, and at least part of T* is newer than T. It is in that case that the reduction of T by means of T* is likely to be an important advance toward the systematization of all empirical knowledge.
>
> T*, or part of T*, may introduce theoretical terms; if so, let us assume that these T*-terms have been introduced by means of the O-vocabulary which was used to introduce the theoretical terms of T. This is possible regardless of the order in which T and T* were proposed. Any term that is either an O-term or a T*-term may be called an O*-term; so at the time T is reduced, the relevant part of our scientific vocabulary is divided into the T-vocabulary and the O*-vocabulary.

Suppose the following O*-sentence is a theorem of T*; we may call it a reduction premise for T: '$\lambda(\theta_1 \ldots \theta_n)$'. The terms $\theta_1 \ldots \theta_n$ are to be names belonging to the O* vocabulary. They may be elementary expressions and belong to the O* vocabulary in their own right, or they may be compound expressions--for instance, definite descriptions--whose ultimate constituents belong to the O* vocabulary. The reduction premise says that T is realized by an n-tuple of entities named, respectively, by the O*-terms $\theta_1 \ldots \theta_n$. Notice that it cannot be true if any of those O* names are denotationless . . . .

If T* yields as theorems a reduction premise for T, and also a suitable set of definitionally expanded bridge laws for T, then T*--without the aid of any other empirical hypothesis--reduces T. For T* definitionally implies the postulate of T, as well as the set of bridge laws. Once T* is accepted, there is no choice whether or not to reduce T. The reduction of T does not need to be justified by considerations of parsimony (or whatever) over and above the considerations of parsimony that led us to accept T* in the first place . . . .

(2) Let T be a theory explaining the regulation of certain biological processes by positing hormones $t_1 \ldots t_n$: chemical substances of unspecified composition, secreted by specified cells under specified conditions and regulating the rates of specified chemical reactions in a specified way. The T-terms $t_1 \ldots t_n$, in this case, purport to name substances. Let T* comprise our body of biochemical knowledge at some later time; T* might imply that certain substances named by chemical formulas $\theta_1 \ldots \theta_n$ realized T, and that they alone did so. To exclude multiple realization of T, T* would have to contain the information that, e.g., a certain gland secretes nothing but the substance with formula θ_1; but we often do have such knowledge.[17]

---

[17] "How to Define Theoretical Terms," Pp. 441-444.

The comments below are meant to apply to a case such as above in which T(r) is taken as following from R. An analogous point can, perhaps, be made when T(r) is introduced otherwise; but I am not sure.

In such a case, let us represent the postulate of the reducing theory R by '$R_1(r)\&T_1(r)$' where $T_1(r)$ represents that part of the postulate from which T(r) is derivable and $R_1(r)$ represents the remainder.[18] For ease of exposition, let us simply write: $R_1(r)\&T(r)$.[19] We then have, given Lewis' definitions: 'r' is synonymous with '$\eta x(R_1x\&Tx)$'.

Now, it might occur, in a given possible world, that there are many things which satisfy T, none that do, or precisely one that does; and similarly with R_1. Let us assume in what follows that there is exactly one thing which satisfies R_1; this is strictly a matter of convenience: it reduces the number of cases to be considered without affecting any of the main points.

If nothing satisfies T then, on Lewis' view, neither t ($=\eta xTx$) nor r ($=\eta x(Tx\&Rx)$) exist. If exactly one thing

---

[18] '$R_1(r)$' could be the null string; I shall assume that it is not.

[19] Technically, we should probably not write it this way; for Lewis does say that 'r', when introduced, is specified as other than the entity which realizes T. Whatever Lewis precisely means by that, he certainly does allow, as we have seen above, postulate of the form: $R_1(r)\&T_1(r)$; that is, postulates of the reducing theory from which T(r) follows. Since my argument would remain precisely the same if I left things as above, I have no qualms about changing from $T_1(r)$ to T(r).

satisfies T, then it is either the same thing which satisfies R_1, or it is not. If it is, then t exists and r exists. If it is not, then t exists but r does not (since there is no one thing uniquely both T and R_1). If there is more than one thing which satisfies T, then if one of those things is that which uniquely satisfies R_1, then t does not exist, but r does. If none of the things which satisfies T satisfy R_1, then both t and r do not exist (all this is, of course, according to Lewis).

So there are circumstances in which t can exist without r, and r exist with t; what is not possible, however, is that r exist and t exist and $t \neq r$. That is, assume we have genes and DNA molecules, hormoes and chemicals (Lewis' example), pains and CFSs. There can not then, in any possible world, be anything other than DNA molecules which are genes, these chemicals which are hormones, or CFSs which are pains.[20] For in any possible world, to be identical with t (pain, such and such a hormone, gene) is to uniquely satisfy t; and if r (CFS, such and such a chemical, DNA) exists, it must also satisfy T, and so be identical to t.

---

[20]What is allowed is that, e.g., hormones of such and such a kind exist when chemicals of such and such a kind do not exist altogether; and vice versa.

Lewis does suggest that sometimes '$r=\iota xTx$' (and not only 'T(r)') follows from the reducing theory. In such a case, we would have the stronger objection that this entailed that if r exists, it must be identical with t.

This argument runs counter to the general thrust of Lewis' argument. It is similar to, though not precisely the same as, the flaw claimed for the materialist position by Putnam and others--chauvinism: e.g., not counting as pain things which clearly are or would be. And it is precisely, among other reasons, to meet this sort of objection, that Lewis has formulated his more sophisticated views; that is (e.g.), to allow that pain might not be CFS.

There remains, however, the equally undesirable consequence here presented: In a situation in which r and t exist, nothing could be r which was not t and nothing could be t which was not r. This poses a serious challenge for Lewis: it shows that the charge of chauvinism still sticks. His appeal to nonrigidity will not extricate him.

Let me conclude by recalling Lewis' reasons for rejecting Carnap's definition of theoretical terms: that it is implicit in the scientist's use of such terms that they are intended to refer to a unique entities. Lewis' program suffers, I believe, from a somewhat similar failing; for it is likewise implicit in the use of such terms that they are about some <u>particular</u> entities--even if it is not determinate on the basis of the O-vocabulary which. Just as having no or multiple realization should falsify a theory, so also should having <u>these</u> (for the relevant "these") entities not behave appropriately do so.

Thus suppose, amazingly enough, that we discover that in a distant part of the (actual) universe, but not locally,

there are little men filling the "functional role" associated with quarks. Lewis would say one of two things in such a situation. Either, (1) there are no quarks and the corresponding theory is false; or (2) quarks are one thing in this region of space and another (little men) in that. But both of these alternatives seem clearly incorrect. There are quarks, but the little men in the distant part of the universe are not. True, if it is part of our theory--and it needn't be--that quarks are the basic elements of the entire universe, then the theory is false until modified. The point, nevertheless, remains: 'quark' refers to one particular kind of thing only; not whatever happens to full the appropriate causal role.

The moral we may glean from our discussion of Lewis is that you can not get your theoretical reductions easy; they must be earned. Even if one believes that, in a sense everything is physical (supervenient upon the physical), and that theoretical terms are functionally definable in such a way that they refer to physical entities only, it does not follow that all special sciences are reducible; nor that all significant theoretical terms refer to physical entities. (Let me point out one last time that this whole discussion is presupposing the existence of universals.) One might accept this as dogma, but there is no argument which has it as its conclusion.

The preceding pages have presented numerous arguments; only a few (and perhaps none) can be claimed to be knock-

down. In most cases, there is a way for Lewis to preserve the integrity of both his views and his formalism. What we have seen is that in doing so, he diverges more and more from traditional physicalist doctrine (when those views countenance, as Lewis does, universals). Whether he diverges far enough so that his views, when fully expanded and its ramifications made clear, will not be welcomed by his brethren, is a question to which there is no determinate answer. On the other hand, if his views are adduced as supporting evidence by traditional physicalists (e.g., Armstrong, Smart), then the objections raised strike true.

KRIPKE

I now turn, in the second half of this thesis, to a consideration of Kripke's antimaterialist arguments. This takes the form of a critical discussion of what I consider to be the most cogent arguments presented against Kripke's views. Only one of the three sections, that dealing with Feldman, deals with arguments aimed directly at Kripke's antimaterialist's arguments. The other two, that of Dummett and Chomsky, deal with Kripke's more general views. Their relevance lies in the fact that if their arguments were sound, they would undermine the premises on which Kripke's antimaterialist arguments are based. Feldman's argument, as we shall see later, is directly relevant to the question of whether Kripke's arguments have force against Lewis.

As already mentioned, since they are so well known, I do not bother to repeat in detail Kripke's views. Let me just, in a few lines, give here the broad general outlines of his antimaterialist arguments. Kripke starts with the Cartesian premise that Descartes could exist without his body, or that Descartes' body could exist without Descartes. Given this, and the necessity of statements of identity containing rigid designators, it follows that Descartes is not identical with his body--for otherwise they would have to be necessarily identical. Kripke also makes similar arguments concerning pains (tokens and types) and brain states. The possible nonidentity of pain states and

brain states, say, can not be explained away. Kripke further argues, by appeal to the notion of epistemic possibility which helps explain away the illusory intuition that heat might not have been molecular motion; for to be in the same epistemic state as pain, is to be in pain; and, consequently, we can not say that what we are imagining is a case in which we are in the same epistemic state as that which we are in when we are in pain, but where we are actually in some other state.

The above outline is intended as a reminder to those who are already familiar with the views. Those who are not, should not expect to gain much from the above. I start in with Dummett's critique of Kripke.

DUMMETT'S CRITIQUE

Dummett's Critique of Kripke*

I wish to discuss Dummett's critique of Kripke's view[1] that proper names could not be synonymous with definite descriptions, or clusters of such descriptions.[2]

Kripke's antisynonymy argument is grounded in two distinctions: That of epistemic vs. metaphysical necessity and that of rigid vs. nonrigid designation. The crux of Dummett's critique is that these distinctions are in reality distinctions of scope; and do not differentiate names from descriptions.

This chapter has the least direct relevance to Kripke's mind/body argument. I have included it because it is widely considered, and correctly so I believe, to be the most forceful argument to date against Kripke's general views.

---

* After completion of this section, a new version of Naming and Necessity was published with a new preface by the author. While he does not go into much detail, it is gratifying to note that what he does say corroborates what I have written. While I have not incorporated this new material into the text, I have made some comments in the footnotes.

[1] In, Frege: Philosophy of Language. All references will be to this.

Dummett's critique of 'The casual theory of reference' and Kripke's Gödel example are not discussed. I have limited myself to what I consider to be Dummett's major criticisms of Kripke's modal arguments.

[2] This view can handle some of the problems the simpler view can not. I shall not go into details since for my purposes, it is not important (See Searle: "Proper Names," Mind 67 (1958): 160-173). For convenience's sake, I shall consider the view that every proper name is synonymous with some one definite description. This simplification does not affect any of the arguments.

The relevance to the mind/body problem it does have is that if Dummett's arguments are sound, then it follows, he contends, that ⌜a=b⌝ does not entail that ⌜necessarily a=b⌝, where a and b are proper names. As we know, that this entailment does hold is a premise of Kripke's antimaterialist argument. Whether he could reformulate these antimaterialist arguments if Dummett's points were conceded, is unclear to me.

I

Kripke's attack on synonymy claims is two pronged.[3] He contends first that the descriptions to which proper names are putatively synonymous do not even standardly "fix the reference" of those names: In many, if not most, cases, the referent of a proper name need not, in the actual world, uniquely satisfy the putatively synonymous description (or cluster thereof). This contention is more or less conceded by Dummett and suffices in most cases for Kripke's antisynonymy claims.[4]

---

[3] I go into the details of Kripke's argument only cursorily. I assume here and throughout that the reader is familiar with the details.

[4] Actually, what Dummett says is that though what Kripke claims may be true, it is not anything Frege would have denied; for Frege did not claim that names generally have the sense of definite descriptions.

Such a reply, whatever its merits as an exegesis of Frege, is certainly disengenuous if intended as a critique of Kripke. The view Kripke ascribes to Frege is one which has previously almost universally been so ascribed. It is a view which has, in any case, received widespread independent support (if it has not indeed, been the dominant view in contemporary Anglo-American philosophy). If nothing else, Kripke's attack, if successful, buries an extremely influential view.

Dummett's exegesis of Frege has received considerable support from a recent article by Tyler Burge, "Sinning Against Frege." Though this article supports Dummett, it also supports my claim that the view Kripke is attacking is the one which historically has been overwhelmingly thought to have been Frege's.

Kripke's argument here will not stand, by the way, if one holds, as does e.g., Katz, that the meaning of a term need not uniquely fix its reference even in the actual world.

Kripke's second attack goes further and says that even if the reference of a name were fixed by some definite description--and he grants there can be such, e.g., 'Jack the Ripper'--the name and description are not synonymous.[5] For (1) the statement 'Necessarily Jack the Ripper is Jack the Ripper' is true on every reading, while (2) 'Necessarily the famous London murderer is the famous London murderer' has a reading on which it is false, and (3) 'Necessarily Jack the Ripper is the famous London murderer' is false on every reading.

A response to this might be: if we are granting that 'The famous London murderer' fixes the reference on 'Jack the Ripper', then there is a reading on which (3) is true: There is a sense in which Jack the Ripper must be the famous London murderer, whoever that may be; similarly with 'St. Anne is the mother of Mary', to use Dummett's more intuitive example (see Fn. 5).[6]

---

[5] The example Dummett uses is 'St. Anne', and 'The mother of Mary'. Since Kripke holds that being the daughter of St. Anne is a necessary property of Mary, using this example might cause some confusion. I shall, therefore, stick to the less intuitive 'Jack the Ripper'. For convenience's sake, I shall use the description, 'The famous London murderer', though it is obviously incomplete as a reference fixer of 'Jack the Ripper'.

[6] It should be emphasized again that we are dealing here strictly with names like 'St. Anne' which have reference-fixing descriptions. It is not the least virtue of Dummett's argument that he actually uses such a name as his example (instead of 'Moses' or 'Aristotle'). For this permits us to see that Kripke's arguments do not have all the intuitive punch they appear to.

At this point, Kripke invokes his epistemic/metaphysical distinction. 'Jack the Ripper is the famous London murderer' may be conceded to be epistemically necessary of (equivalently) a priori; i.e., in any world in which we fix the reference of 'Jack the Ripper' as we do in the actual world, we can truthfully and a priori assert: Jack the Ripper is the famous London murderer.

But there being an epistemic reading on which (3) is true does not conflict with Kripke's above claims, for (1), (2), and (3) are concerned strictly with what Kripke calls metaphysical necessity; and (3) is not metaphysically necessary. Jack the Ripper might have channeled his drives into more conventional outlets; and a statement is metaphysically necessary only if there are no circumstances under which it would be false.

(1), (2), and (3) are intended to show the distinct behavior of names and descriptions in modal contexts. But, replies Dummett, to discern such distinctions, one must have antecedently committed one's self to Kripke's metaphysical/

(Fn. 6 continued)

Thus, Kripke argues that even if we grant, counterfactually, that, 'The leader of the Israelites' fixes the reference of 'Moses', 'Moses was the leader of the Israelites' is still contingent--Moses might have remained his whole life in Pharaoh's court. But much of the intuitive force of this argument is deceptive; it comes from our inability to put ourselves in the appropriate counterfactual frame of mind (i.e., with 'The leader of the Israelites' fixing the reference of 'Moses'). This is confirmed by the fact that the analogous argument in the case of 'St. Anne might not have been the mother of Mary' is (though I still think valid) considerably less convincing.

epistemic distinction; otherwise, there are no differences to be discerned between the behavior of names and descriptions. Thus, (2) has a reading on which it is false and one on which it is true, and (3) likewise has a reading on which it is false and one on which it is true (Kripke's epistemic reading). The reading on which (2) and (3) are both false, contends Dummett, are those on which we take the description and name, respectively, outside the scope of the modal operator--i.e., the de re reading. We might paraphrase these readings as: The famous London murderer is (has the property of being) necessarily the famous London murderer; and: Jack the Ripper is (has the property of being) the famous London murderer. While the readings on which they are true are those with the description and name, respectively, within the scope of the modal operator--i.e., the de dicto reading. These may be paraphrased as: necessarily, the famous London murderer is the famous London murderer; and necessarily, Jack the Ripper is the famous London murderer.[7] So that Kripke's distinction is really one of scopes.

Let us turn now to the rigid/nonrigid distinction. The difference in truth value between (1) and (2) (on one reading) is due, according to Kripke, to the fact that 'Jack the Ripper' is, and 'The famous London murderer' is not,

---

[7] I count as de dicto any reading of a statement with a modal operator to the left, and nothing to the left of it (and with no free variable); this is so whether what is construed as having necessity attributed to it is a linguistic entity or otherwise. I shall assume throughout that there is only one modal operator in any statement under consideration.

a rigid designator. Indeed, Kripke's criterion for a term being a rigid designator is that there be no sense (on any reading of scopes) on which it is true to say ⌜α might not have been α⌝. His other criterion for rigidity is that a term is rigid if it refers to the same individual in every possible world in which that individual exists.[8] This rigid/nonrigid distinction is what Kripke takes to be the crucial semantic difference between names and descriptions.[9]

Dummett's analysis of this distinction is that it too is just the distinction of scopes in sheep's clothing. To say that a term is rigid is just to say that we take it as occurring (in a particular sentence) out of the scope of the modal operator, and to say that it is nonrigid is to say that we take it as occurring within the scope of the

---

[8] Thus, 'Jack the Ripper' refers to the same things in every possible world, while 'The famous London murderer' does not; there will then be conditions under which the two terms are not coreferential and hence, conditions under which 'Jack the Ripper is the famous London murderer' is false. Thus, Dummett puts it (more formally):

> If we say that 'The teacher of Alexander might not have taught Alexander', we may represent this as: $\Diamond\neg T(\iota x;Tx)$ and this might be expanded using Russell's theory of descriptions as either (1) or (2)
>
> (1) $\Diamond\exists y\,[\forall x(Tx \leftrightarrow x=y) \wedge \neg Ty]$
>
> (2) $\exists y\,[\forall x(Tx \leftrightarrow x=y) \wedge \Diamond\neg Ty]$

[9] Perhaps it would be better to talk, even on Kripke's view, of rigid uses of designators. At a minimum, Kripke believes that the classes of terms he calls rigid (e.g., proper names, demonstratives) are standardly used rigidly.

modal operator. Since both descriptions and names can occur within or without the scope of a modal operator, the rigid/nonrigid distinction does not differentiate names from descriptions.

Before turning to the actual details, there is a possible confusion which should be cleared up. At first sight, it seems that Dummett's analysis of the epistemic/metaphysical distinction is inconsistent with his analysis of the rigid/nonrigid distinction. For on his view, for a term to be rigid is for it to occur outside the scope of a modal operator, while for an assertion to be epistemic, is for it to have its referring terms within the scope. There could not then be, it would appear, any statements of epistemic necessity with a referring term which is a rigid designator; for that term would then have to be both within and outside the scope of the modal operator. Yet, we may, with perfect reasonableness, contend that there can be such assertions: e.g., 'Jimmy Carter might be a robot in disguise'. Nor, for the same reason, would it seem there could be an assertion of metaphysical necessity with a referring term which is a nonrigid designator: e.g., 'The first man on the moon might have had red hair'.

And in fact, Dummett is perfectly happy (indeed insists) that one can have such assertions. The apparent contradiction is eliminated by a more precise statement of Dummett's view (though he does not put it in these terms): To use a term

rigidly is to have it out of the scope of a modal operator, and to use it nonrigidly is to use it within the scope of a modal operator; and there is no special class of referring terms which is intrinsically used rigidly or nonrigidly: Names and descriptions can both be used within or without the scope of a modal operator--i.e., rigidly and nonrigidly. Thus, an assertion of epistemic necessity containing a proper name will just be a case of an assertion with a term--which is perhaps most often used rigidly--being used nonrigidly. That this is possible is, in a sense, the point of Dummett's argument.

The above gives the crux of Dummett's argument: Kripke's rigid/nonrigid distinction is really one of scopes. Therefore, the most that distinction can show is that the conventions of our language have it that proper names, in modal contexts, are standardly interpreted as being outside the scope of modal operators, and definite descriptions within. But even this is not so, since both proper names and definite descriptions can occur within or outside modal operators. (Dummett does make one concession to Kripke which we shall discuss later.)

My evaluation of these arguments will procede in three subsequent sections. Section II will discuss whether or not Dummett's arguments have enough force to decisively defeat Kripke; Section III will discuss whether there are any effective counterarguments Kripke has available; and Section IV will reconsider some of Dummett's arguments from a slightly different perspective.

II

This section will proceed in the following manner. First, I will present a certain assumption of Kripke's arguments. Then I will ask of Dummett's objections whether (1) they show this assumption to be wrong, or (2) they show that Kripke's conclusions do not follow from this assumption.

My conclusion shall be that Dummett shows neither. That he has, essentially, made a different assumption than Kripke's; and that, consequently, unless further argument is brought for or against either view, it must presently be construed as a stalemate.

The assumption of Kripke's I am referring to is as follows: (1) The way we evaluate the truth conditions of a de dicto assertion of metaphysical necessity is something like this: We take the sentence in question, and each possible world, and determine whether the sentence, with the sense it presently has, is true to the facts (in the Tarski sense) in each possible world.[10] The important point here being that it is evaluated as a de dicto and not de re assertion.

---

[10] Thus, in his new introduction, Kripke says the following ('(1)' refers, in this and the next footnote, to the sentence, 'Aristotle was fond of dogs'):

> Presumably, everyone agrees that there is a certain man--the philosoper we call 'Aristotle'--such that, as a matter of fact, (1) is true if and only if he was fond of dogs. The thesis of rigid designation

(2) Rigid designators are to be distinguished from nonrigid designators by their behavior in de dicto metaphysical contexts. Thus, Kripke says:

> Most of the things commonly attributed to Aristotle are things that Aristotle might not have done at all. In a situation in which he didn't do them, we would describe that as a situation in which Aristitle didn't do them. This is not a distinction of scopes (emphasis mine), as happens sometimes in the case of descriptions, where someone might say that the man who taught Alexander didn't teach Alexander.
>
> This is Russell's distinction of scopes. (I won't go into it.) It seems to me that this is not the case here. Not only is it true of the man Aristotle that he might not have gone into pedagogy; it is also true that we use the term 'Aristotle' in such a way that, in thinking of a counterfactual situation in which Aristotle didn't go into any of the fields and do any of the achievements we commonly attribute to him, still we would say that was a situation in which Aristotle did not do these things.[11]

(Fn. 10 continued)

> is simply--subtle points aside--that the same paradigm applies to the truth conditions of (1) as it describes counterfactual situations. That is, (1) truly describes a counterfactual situation if and only if the same aforementioned man would have been fond of dogs, had the situation obtained.

[11] Though my reading of this quote is not the only possible one, it is the most reasonable, I feel--especially when taken in context. Also, notice the following quote from Kripke's preface; in it, the contrast he is referring to is that between rigid and nonrigid designators, and sentence (1) is 'Aristotle was fond of dogs'.

> My point, however, was that the contrast would hold if all the sentences involved were explicitly construed with small scopes (perhaps by inserting a colon after 'that'). Further, I gave examples (referred to above) to indicate that the situation with names was not in fact parallel to that with large scope descriptions. Proponents of the contrary view often seem to have overlooked these examples, but this is not my point here. The contrary ..... what about:

Let us turn now to the details of Dummett's arguments. First, with regard to the metaphysical/epistemic distinction, Dummett advances two arguments. The first is an argument from economy: It is conceptually simpler to advert to scope distinctions only, to explain the modal ambiguities evidenced, than to have to advert to Kripke's epistemic/metaphysical distinction in addition.

Given that we are assuming that metaphysical necessity (as well as epistemic), as construed by Kripke, is to be evaluated as de dicto necessity, if Kripke's metaphysical/epistemic distinction is prima facie plausible, then Dummett's argument from economy carries no weight in and of itself. One can hardly eliminate a clearly perceived distinction by saying that it would be simpler if it were not so perceived.

The question then is, "Is the distinction prima facie plausible?" Consider again, 'Jack the Ripper is the famous London murderer'. Is there, as Kripke would have it, a clear nonepistemic de dicto sense on which this statement is possibly false? Yes: there are circumstances in which this statement would not be true, to wit, Jack the Ripper goes into academia. STOP. There is no call to go further and talk of essential properties or de re modality. Kripke's

---

(Fn. 11 continued)

(4) What (1) expresses might have been the case. Doesn't this express the desired assertion, with no scope ambiguities? If not, what would do so?

See also Fn. 10.

notion of metaphysical de dicto modality hinges on whether there could be circumstances under which the given statement is true (or false). And that is all he commits himself to.

Dummett's second and primary argument purports to show that Kripke's notion of metaphysical modality reduces to that of de re modality; consequently, there is once again no distinction to be discerned between names and descriptions in modal contexts. Thus (I quote in full);

> What then is the fact whose contingency we express by saying that the standard metre rod might have been shorter or longer than 1 metre, but which is not expressed when we say a priori that it is 1 metre long, or that it has the length it has? So long as we pose the question this way there does not seem to be any satisfactory answer. Rather, it is not so much that some contingent fact obtains, at least, if we understand a fact as something that can be expressed by means of a sentence understood in some specific sense, but that a certain object, namely the standard metre rod possesses a contingent property, that of being 1 metre long; or perhaps that a certain length, namely a metre, possesses the contingent property of being the length of standard metre rod. If we refer to the rod as 'the standard metre rod' then we guarantee that (provided we are referring to anything at all) we are referring to a length which is that of the standard metre rod. But that very rod which we refer to might have been of a different length; that very length which we refer to might not have been that of the standard metre. This sort of contingency cannot be grasped in terms of the notion of a contingent fact, but only in terms of that of an accidental property. And, indeed, Kripke himself strongly emphasizes the importance for his doctrine of the distinction between essential and accidental properties of an object. But what this means is that we cannot attain to the required notion of contingency by concentrating on the linguistic form: 'It is contingent (possible, necessary) that.....', where the gap is to be filled by an entire sentence; we have, instead to understand the form: ' It is contingently (possibly, necessarily) true of {that ..... {.....'. We have to explain, not what it is for the sentence, 'The standard metre rod is 1 metre long', to

> be contingently true, but what it is for the predicate 'ξ is 1 metre long' to be contingently true of an object; equivalently, we have to understand not the sentence '$\Diamond$(the standard metre rod is not one metre long)', but the predicate '$\Diamond$(ξ is not 1 metre long)'. Just the same will be the case with the contingency which we express by saying that St. Anne might not have been a parent. We can not understand this as relating directly to the status of the sentence 'St. Anne might not have been a parent', as this might be used to express something known a priori, but as saying of St. Anne, that she possessed the accidental property of being a parent. 'St. Anne might not have been a parent' should not be rendered as 'It is possible that St. Anne was not a parent' but as 'It is true of St. Anne that she was possibly-not-a-parent'. But what this means is that, in order to understand the sort of contingency Kripke alleges to exist in these cases, we are compelled after all to invoke just the notion of scope to which Kripke appealed in the case of definite descriptions. In 'St. Anne might not have been a parent', the name 'St. Anne' must be construed as being within the scope of the modal operator: this is what is implicit in Kripke's account in terms of accidental and essential properties, as against contingent and necessary facts or statements. It is thus not merely that the uniform explanation, in terms of scope, of the ambiguity that occurs when either definite descriptions or proper names occur in modal contexts is preferable, because more economical, than having, in the latter case, to introduce the a priori/necessary distinction, it is that, in order to understand the notions of necessity and contingency that Kripke uses, we find ourselves forced to appeal to the notion of scope, for proper names as well as definite descriptions (pp. 124-125).

Let us consider first Dummett's puzzlement over Kripkean "facts": 'What then is this fact whose contingency we express by saying that the standard meter rod might have been shorter or longer than one meter? There does not seem to be any satisfactory answer, if we understand a fact as something that can be expressed by means of a sentence understood in some specific sense.' This seems hardly a fair query. We might just as legitimately ask of Dummett: 'What

then is this fact which can be expressed by means of a sentence understood in some specific sense?' Indeed for Kripke, it is the same fact (actually Kripke would say: same statement; see below) which is said to be contingent as that which is known a priori. If there is, as Dummett thinks, a clear sense to the latter, then there is for Kripke a clear sense to the former and the mystery of facts disappears.

Actually I am here being somewhat unfair to Dummett. Immediately preceding this quote, he has argued that it can not be the same fact which is contingent and a priori. I do not think his argument is convincing.[12]

---

[12]Dummett's argument is basically by way of example. Here is one. Suppose someone in 1001 A.D.--at which time it was true by stipulation that Christ was born in 1 A.D.--asserts that Christ was born in 1 A.D. That, says Dummett (and let us agree) is a paradigm of the type of statement which is claimed by Kripke to be contingent a priori. But there is here, claims Dummett, no one fact expressed which is both a priori and contingent.

Thus, suppose the person making the assertion has no idea of what year it is. His claim that Christ was born in 1 A.D. is indeed a priori. In a sense, his knowledge is of the conventions of the dating system. But there is nothing known here which is contingent. Perhaps if he knew further that it was 1001, he would know the contingent fact that Christ was born 1000 years ago. But this further fact is not known a priori since it is inferred from other contingent knowledge.

Analogously, if he knew just that Christ was born 1000 years ago (and does not know the basis of the dating system), all he knows is some contingent fact and nothing a priori. So that Kripke's claim that we have one thing, being a priori and contingent, is based on an ambiguity deriving from the different possible ways of fixing the reference of '1 A.D.'

Though this argument is appealing, it is not (obviously anyway) correct. The supposition that the person making the assertion knows only the dating system (and no other related fact including the present date) or only that Christ

Even if it were correct, all it would show is that one and the same fact could not simultaneously be contingent and a priori. It would not show that Kripke's distinction between epistemic and metaphysical necessity applied to different sorts of facts; that one case but not the other must be analyzed in terms of necessary properties. Thus, Dummett himself allows the feasibility of "ontic" necessity (about which more later) versus epistemic. On his interpretation, there can be no a priori (epistemically necessary), (ontically) contingent statements. Yet, various statements

---

Fn. 12 continued)

was born 1000 years ago, is not, I believe, coherent. In any case, the situation, even if feasible, seems to me to be completely symmetrical; and I do not see how we can decide that in one case, we have a priori knowledge, while in the other, knowledge of a contingent fact.

Thus, if all that a person knows is Christ was born 1000 years ago, there is a sense in which he does not know what year it is now: he knows it is this year (1001), but not that this year is 1+ however many years have passed since the birth of Christ. Similarly, if all one knows is the calendar convention, there is sense in which one doesn't know what year Christ was born in: one knows Christ was born in whatever year Christ was born in (1 A.D.), but not that he was born 1000 years ago. These two cases seem completely equivalent as far as the having or lacking of a prioricity (or contingency) goes.

It seems thus that Dummett's idea of breaking up the individual's knowledge into an a priori and a contingent component is not satisfactory; and our knowledge does not, in fact, consist in such independent facts. There is thus room for our statements to be a priori contingent without there being any ambiguity. This question is intricate and more could be said; it seems clear enough though that Dummett has not yet shown that a statement could not be a priori contingent.

can, univocally, be said to be either ontically or epistemically necessary or contingent. Similarly here, that a statement can not be a priori contingent does not, in itself, entail that the two sorts of necessity apply to different sorts of statements or facts.

Actually, the whole question of "facts" is something of a red herring. Kripke standardly discusses the necessity of statements and he explicitly makes the point of distinguishing the metaphysical necessity of statements (and not facts) from de re necessity. How then is Dummett able to conclude, "We can not attain the required notion of contingency by concentrating on the linguistic form 'It is contingent (necessary, possible) that .....' where the gap is to be filled by a whole sentence?" That, according to Kripke, is precisely the required notion; it is the assumption we are working with. Have any reasons been adduced to deny that assumption?

One thing which we should clear up right away is Dummett's apparent suggestion that Kripke appeals to the notion of necessary properties in his elucidation of the epistemic/metaphysical distinction. (I.e., "And, indeed, Kripke himself strongly emphasizes the importance for his doctrine of the distinction between essential and accidental properties of an object.") It is certainly true that Kripke defends the coherence of essential attribution. It is equally true that that notion is <u>not</u> appealed to in his

analysis of metaphysical necessity: Indeed, this is our assumption, and it is an assumption pretty well born out by the text (from which I've given some quotes above). Dummett will have to produce an independent argument if he is to convince us of this point.

Dummett's reasoning, as far as I can tell, appears to be something like this. Facts, for Kripke, consist in objects having certain properties. To assert of any such fact that it is necessary is, therefore, to appeal to the notion of necessary property; and that is to take the assertion of necessity on its de re reading.

Now, even if we do talk of facts, how does it follow that if one is committed to necessary facts, then one is committed to necessary properties? There appears nothing inconsistent or unreasonable in believing that an object's having a certain property (i.e., a fact) is necessary--independently of how the object in question is referred to--without committing one's self to the existence of necessary properties, or just plain properties for that matter.

And even if one were so committed to necessary properties, there is still room for the existence of (distinct) necessary facts (or better statements): and Kripke need merely have his argument be concerned with these.

We have seen that nothing Dummett has said counts against Kripke's assumption about the analysis of metaphysical necessity, or the validity of the conclusions drawn from that assumption (except, perhaps, the weak claim

that it is simpler to just advert to scope distinctions).

But, of course, neither have I given any positive reasons for denying that appeal to scope distinctions is the appropriate method of analyzing the data. As far as anything has been said, there really seems to be two separate analyses with no strong reason to support one over the other. And Dummett himself says things in this vein:

> His wish to dispense with the notion of sense for proper names leads him to regard a fact as consisting, e.g., in the possession by an object of a certain property, or in two objects' standing to one another in a certain relation. A fact, so conceived, may be taken as forming the content of a particular statement, but it certainly cannot be identified with the thought expressed by the statement, as Frege conceives of it, and hence cannot properly speaking be said tc be an object of knowledge at all. The knowledge which someone expresses by means of an assertion (when it is knowledge) is the knowledge that the thought expressed by the sentence used to make the assertion is true; it cannot, properly speaking, be taken to be the knowledge that the fact obtains (in Kripke's sense of 'fact') which is the content of the assertion. Thus, for instance, Kripke's notion of facts leads straight to the conclusion, willingly drawn by Kripke, that the fact which is the content of a true statement of identity is always a necessary one: for it is just the fact that a certain object bears to itself that relation which every object bears to itself and to no other (p. 126).

We have witnessed no reason to think one analysis is better than the other.

As long as I have quoted the above, let me jump ahead a little. My last section will be concerned with a possible line of argument on Dummett's behalf: That even though a de dicto metaphysical/epistemic distinction may exist, it is solely the latter which is relevant to a term's meaning.

There are suggestions to this effect throughout Dummett's discussion; and the above might also be so construed.

Let us turn now to the rigid/nonrigid distinction. Dummett's claim, let us recall, is that this is in reality a distinction of scopes. And that both proper names and definite descriptions can be given either large or small scope. Since Dummett's style usually leaves him open to multiple interpretations, I shall, once more, quote in full.

> The thesis that proper names are rigid designators is expressed in terms of the metaphor of possible worlds, and hence, to give it substance we must remove the metaphor. And, as soon as we try to do this, we see that it concerns nothing other than our old notion of the scope of a term in a modal context. For definite descriptions, there is no distinction between their meaning and the way their reference is determined; in other words, the way in which the reference is determined in the real world is carried over into each particular possible world. In any possible world, the referent of 'The man who led the Jewish people out of Egypt' is the one and only object (if any) which, in that world, satisfies the predicate 'is a man who led the Jewish people out of Egypt'. But the whole point of saying that, for a proper name, its meaning diverges from the way in which its reference is determined is to make clear that the latter is not taken as carrying over into whatever possible world we are concerned with. So that there is a possible world in which St. Anne had no children, involves that, in that world, the reference of the name 'St. Anne' is not determined in the way in which it is determined in the real world; rather, it is determined in the world via its reference to the real world, i.e., as being to the same woman as the one who is its referent in the real world. This can only be interpreted as the thesis that, in a modal context, a definite description must always be construed as lying within the scope of the modal operator, while a proper name must always be construed as lying outside its scope. To assign a term a reference varying from one possible world to another is just to take it as having, in each possible world, the refernece which it would have in that world; coversely, to assign it a constant reference is to take it as having, in each world,

> just that reference which it has in the real world. But to take a term in the former of these two ways is precisely to treat it as being within the scope of the modal operator, while to take it the second way is to treat it as falling outside that scope. Thus, when Kripke says that it would not be true that: the teacher of Alexander didn't teach Alexander, he is intending to convey that within any possible world, it would never be true to say that 'The teacher of Alexander didn't teach Alexander'. Here the definite description is taken to have as referent, within each possible world, the unique object (if any) which in that world satisfies the predicate 'taught Alexander'; and we display our adoption of this interpretation by rendering the sentence with the description taken as falling within the scope of the modal operator, namely as: $\neg\Diamond\exists y \forall x [(Tx \leftrightarrow x=y) \cdot \neg Ty]$. When, however, we assert that the teacher of Alexander might not have taught Alexander, we are treating the definite description as having as its constant referent, that referent which it has in the real world, and this amounts to taking the description to lie outside the scope of the modal operator.
>
> Kripke's doctrine that proper names are rigid designators and definite descriptions nonrigid ones thus reduces to the claim that, within a modal context, the scope of definite descriptions would always be taken to exclude the modal operator, whereas the scope of a proper name should always be taken to include it. Even if this were so, it would not demonstrate the non-equivalence of a proper name with a definite description in any very strong sense; it would simply show that they behave differently with respect to ad hoc conventions employed by us for determining scope (pp. 127-128).

The issues here are not really very different from those encountered immediately above. This should come as no surprise. If Dummett is correct in his claim that the only coherent (or perhaps as we shall discuss in Section IV, relevant) de dicto notion of necessity is the epistemic, there are no grounds to distinguish names from descriptions. (Even Kripke would admit this.) For Dummett, both claiming

that an assertion is epistemically necessary and that a term is nonrigid, is the same as saying that the relevant terms fall within the scope of the modal operator; and that is equivalent to the claim that the reference of the term in other possible worlds is determined by what the term would refer to in each such world. Equivalent remarks apply to his analysis of rigid terms.

If, on the other hand, Kripke is correct, and we can make out the metaphysical de dicto reading, then the (nonscope) distinction between names and descriptions is straightforward, one has only to consider the statements, "Aristotle might not have been the teacher of Alexander," and "The teacher of Alexander might not have been the teacher of Alexander" construed (metaphysically) de dicto (See also footnotes 10 and 11 on this).

So, once again, the issues seem to depend on a different assumption on how to construe the truth conditions of modal statements. I've already given what I take to be Kripke's. Perhaps we can say that Dummett's view is that a (modal) statement's truth conditions is determined by its "logical form"; that modal assertions allow for two such (de dicto and de re), each of which is applicable equally to sentences containing proper names or descriptions.

Both Kripke and Dummett's views follow naturally from their assumptions. But, so far, no strong reason has been adduced to support one over the other.

I turn now to a problem first encountered in Section I. Names can be distinguished from descriptions (reference fixing or otherwise) by virtue of the fact that we can truthfully say, e.g., 'The teacher of Alexander might not have been the teacher of Alexander' while we cannot say, 'Aristotle might not have been Aristotle'. On Kripke's criterion, 'Aristotle' is rigid while 'The teacher of Alexander' is not.

In response to this, Dummett makes a limited concession to Kripke. "The grain of truth in Kripke's view in the behavior of proper names after verbs like 'be'." That this is so has to do not, claims Dummett, with the general behavior of terms in modal contexts, but rather with the fact that, "we do not regard such a predicate as 'α is St. Anne' as standing for a property that can be acquired." (While 'α is the mother of Mary' would, presumably, stand for such a property.)

One might think such a concession all the concession Kripke needs; not so, avers Dummett. What is involved in this analysis is really (a cluster of) two descriptions synonymous with the proper name: "It is not exactly accurate to say the name 'St. Anne' has the sense of 'The mother of Mary': Rather it has a sense such that it is replaceable either by 'The mother of Mary' or by 'The woman who was to be the mother of Mary', according to context."

As another example, let us take 'Jack the Ripper', and see if this move works. For it to do so, Dummett must have it that an appropriate replacement of both occurrences of 'Jack the Ripper' in 'Jack the Ripper might not have been Jack the Ripper' will yield a statement which has, as does the original, no true reading. Replacing both occurrences of the name by the same description will clearly not do. What Dummett intends is that we replace the first occurrence by 'The famous London murderer' and the second by 'The man who was to become the famous London murderer'. The outcome is a sentence, 'The famous London murderer might not have been the man who was to become the famous London murderer,' which, on Dummett's analysis of the modalities, has no true reading.[13]

But the above does not yield the desired conclusion: i.e., that there is a sentence (containing only the appropriate descriptions) which has no (tout court) true reading; Dummett has only shown there is no true reading on his conceptions of modality. There may be true readings on other conceptions of modality. Indeed one such is Kripke's: The famous London murderer might not have been

---

[13]Dummett defines an essential property as one which is, during every time of an object's existence, a 'presently essential' property. A presently essential property is one which, once an object has it, it can not lose and remain the same object. Being a man who was to be the famous London murderer would, Dummett would have it, be an essential property, and 'The famous London murderer might (in the Dummett sense) not have been the one who was to be the famous London murderer' is false on all readings.

the famous London murderer; in which case he likewise would not have been the one who was to become the famous London murderer. Hence, there is a sense in which 'The famous London murderer might not have been the one who was to become the famous London murder' is true. Unless Dummett shows that there is some incoherency in Kripke's conception of modality, and he has not, his argument here too will not stand. And once again, the issue seems to depend on differing conceptions of how to analyze the modalities.

III

The conclusion of the last section was that we had reached a stalemate. In the present section, I wish to see if any positive arguments can be marshalled on Kripke's behalf.

The first point is one that has already been touched upon. Both Kripke and Dummett agree on the existence of a de dicto epistemic reading of modal assertions. The bone of contention is, as we have seen, whether or not there is a further de dicto metaphysical reading of modal statements. There does seem to be a fairly nonproblematic sense to this. A sentence is true if it corresponds to the facts (again, we can unpack this à la Tarski). It is necessarily true, if it corresponds in a similar manner to the facts as they obtain in each world (we can think of a world as all the facts there are). This seems to me relatively straightforward, and unproblematic, Dummett's attempted denogration of the concept not withstanding.

Nor, as already mentioned, is there really any reason to think that such appeals to facts on other possible worlds invokes the notion of necessary properties. Thus, consider someone who believes that a statement is true if it corresponds to the facts: and hence believes in facts. It would be wrong to claim that merely by adherence to this theory of truth he is committed to hypostatizing properties; i.e., that since facts are nothing but objects having properties to say that 'A is P' corresponds to the

facts is to say that there exists a property P had by A. This reasoning is clearly invalid; one can be committed to facts without recognizing properties. My point above is similar; the facts are just on other possible worlds. Or consider one who chooses to interpret 'necessarily' as 'compatible with the laws of nature'. Can he not make assertions of de dicto metaphysical necessity (which can still be distinguished from epistemic) without committing himself to 'compatible with the laws of nature properties'?

Indeed, just this sort of Tarski-type analysis extended to other possible worlds has been used by H. Field (in his previously cited article) for quite independent reasons. Thus he writes:

> For instance, for a sentence of form ⌜P(b)⌝ where P is a predicate and b a name, the definition will read
>
> > ⌜P(b)⌝ is true at w if and only if there is an object x that b denotes (in the actual world) and a property Z that P stands for (in the actual world), and w is a world in which x exists and has Z.
>
> For example, 'Bertrand Russell is hairless' is true at any possible world w in which Bertrand Russell (i.e., the person denoted in the actual world by the name 'Bertrand Russell') exists and is hairless (i.e., has the property that the word 'hairless' stands for in the actual world).

Given such an analysis, it is straightforward how we would analyze the (de dicto) notion of metaphysical necessity, and rigid designation. (Notice there is no reference to necessary properties.) I think there are problems with this treatment (e.g., its ability to deal with

definite descriptions instead of names). What is pertinent though is that Field exhibits no need whatsoever to defend the coherency of his formulation. Further, if materialism is true, and a conception as above is necessary for an adequate materialist account of the notion of truth, we would have proof positive of the notion's coherency.

Besides its prima facie plausibility, there are examples we can bring on Kripke's analysis' behalf: 'Cats might be robots' I say, to use the venerable example. This is metaphysically false, read both the de re and de dicto; there are no possible circumstances under which cats (the mammalians we know and love) could be robots. But it can be given an epistemic interpretation on which it is true: I could be in the same epistemic state as I am now and it turn out that the things I've been calling cats are really robots in disguise. On the other hand, if off in the distance I vaguely discern an object which, though I don't know it, is a table, I may say 'The object in the corner might be an elephant', which seems true read de dicto: epistemically or metaphysically; there are circumstances under which the statement would be true. But false read de re.[14]

---

[14]It might be objected to the validity of this example because of the indexicality of 'The object in the corner'. But why should that make a difference? Below I discuss 'I' as a rigid designator and its appropriateness might there too be queried since it can apply to many people in any given possible world. But if this is a problem, it is a problem faced by almost all rigid designators including the paradigms: proper names and terms like 'this' and 'that'. The crucial

These examples should not deceive us into being overconfident; they do not, by themselves, prove that Kripke's distinction exists; Dummett will say that in the former case he can distinguish only one reading on which the sentence is false, while in the latter only one on which it is true; he can never be forced to concede Kripke's distinction by such examples, for any perceived differences in modal assertions can always be construed as either a difference in scope (if there is a difference in truth value) or not to exist at all (if there is no difference in truth value). Nevertheless, these examples do add plausibility to Kripke's case.

(Fn. 14 continued)

property of rigid designators which make them such is their demonstrativeness; as is clear from Kripke's exposition (consider again 'this' and 'that'). On that count, 'I' certainly qualifies. Further, note the following from Kripke's new Introduction:

> In practice, it is usual to suppose that what is meant in a particular use of a sentence is understood from the context. In the present instance, that context made it clear that it was the conventional use of 'Aristotle' for the great philosopher that was in question. Then given this fixed understanding of (1), the question of rigidity is: Is the correctness of (1), thus understood, determined with respect to each counterfactual situation by whether a certain single person would have liked dogs (had the situation obtained)?..... To speak of 'the truth conditions' of a sentence such as (1), it must be taken to express a single proposition--otherwise its truth conditions even with respect to the actual world are indeterminate. Thus, ambiguous words or homonyms (perhaps 'dog' is (1)) must be read in a determinate way (canine!), indexicals must be assigned determinate references, syntactic ambiguities must be resolved, and it must be fixed whether 'Aristotle' names the philosopher or the shipping magnate. Only given such a reading can Russell propose an analysis such as (3)--rightly, no one ever faulted him on this score.

Finally, there is Kripke's example (quoted already in Fn. 11) which is not so obviously answerable by Dummett. I have changed the example so as to make it as favorable as possible to Dummett--and to show that Kripke still makes his point. Consider the following:

(1) St. Anne was the mother of Mary.

(2) What (1) expresses might not have been the case.

(2) appears true; on one reading, anyway. Yet, it can not plausibly be construed as making a de re assertion of modality (after all, the object of the modality is what (1) expresses). Nor can it be taken to be attributing epistemic modality; for on <u>that</u> reading, both Kripke and Dummett agree (2) is false. So its true reading must be the metaphysical de dicto one.

Finally, there is the following: We have been assuming that according to Dummett, the only reading of de dicto necessity is what Kripke calls the epistemic. We suggested that Dummett might hold (more or less) that truth conditions for modal statements are dependent on their "logical form"; dependent, that is, solely on how we read scope. If, however Dummett were to admit as valid some other nonepistemic de dicto interpretation, then his position would be much weaker. For then, no longer would it just be a question of two different ways of evaluating truth conditions: For Dummett would admit nonepistemic de dicto readings: It would be incumbent upon him to say what is wrong with Kripke's.

But this is what Dummett concedes. He admits the feasibility of a "realistic" notion of (de dicto) necessity. Such an interpretation would hold that "the sense of our sentences is given in such a way as to relate to their determination as true or as false, by a reality existing independently of us, and that, in a well constructed language, every sentence will thus be determined true or false, independently of our capacity, even in principle, for recognizing what truth value it has" (sounds close to Kripke's metaphysical necessity, does it not).

Dummett proceeds to give an explicit elucidation of what such a notion might be like, which he calls 'ontic' necessity:

> If for the understanding of a given sentence it is necessary to invoke the conception of a being whose powers of observation or mental capacities transcend ours in a given respect, then the statement, if true, is ontically necessary, if it would be epistemically necessary for such a being--i.e.: if it could be known a priori by him.

Now a few comments:

(1) This criterion itself involves reference to nonepistemic necessity--i.e., the 'would' and 'could'.[15]

---

[15] If the notion of ontic necessity as so explicated is legitimate, then so must Kripke's notion of metaphysical necessity be--for the former adverts to the latter. It could not be that the 'could' is to be taken relative to our understanding--i.e., epistemically--for then we would have to understand the nature of his a priori knowledge of the sentence and we would therefore be able to understand it a priori ourselves, contra hypothesis. Nor would the 'could' be taken relative to the being's understanding, because then we would have to posit yet a higher being with whom the same problem would arise.

(2) Let us make a slight variation on Dummett's criterion and take as the higher being, God. Let us consider Him pondering which of all the possible worlds is best so that He may bring it into existence. In particular, He ponders a world in which "Moses led the Israelites out of Egypt," and "Jack the Ripper is the famous London murderer" is false. These assertions, assuming as we are that the descriptions fit the reference of the names, are epistemically necessary. Yet, by a slight variation of Dummett's criterion, they come out ontically (read: metaphysically?) contingent, so this would confirm Kripke's interpretation.

(3) There is no connection made between a "realistic" interpretation, as first described, and the "ontic" one just presented. No reason is given why a realistic interpretation has to be this one and no other.

(4) Which brings us back to our original point. There is no reason why, if we are granting the coherency of "realistic" interpretations, that we can not take Kripke's to be one such. We have seen previously Dummett's attempt to discredit that notion--i.e., by claiming it was nothing more than essential attribution; and we saw that he could not make the charge stick. Given the assumptions we were working under, we declared a draw; but from the present perspective, it appears that Dummett remains with a debt to discharge: showing something wrong with Kripke's notion of metaphysical necessity.

Dummett does have two things to say about this, when discussing ontic necessity. First:

> None of this has, however, very much to do with the topic of the behaviour of proper names and definite descriptions in modal contexts. This can be seen from the fact that both the notions of ontic and of epistemic necessity that have been discussed concern the status of whole sentences. They, therefore, can be used, without supplementation, only to explain the occurrence of initial modal operators, whereas Kripke's account of definite descriptions involved treating such operators as capable of standing within the scope of other operators, in particular of quantifiers. In order to provide a sense for modal operators in such contexts, we have to take a step of quite a different kind from that of distinguishing ontic from epistemic necessity: we have, namely, to explain when a predicate containing a modal operator, for instance the predicate '{ might not have been a leader' or '{ might not have been 1 metre long', is true of an object.

But this is really begging the question. When we construed Dummett's claim to be that the only issue involved was "logical form," and the only coherent sort of necessity was epistemic, the above could be plausibly argued, given those assumptions.

But now given that we are accepting the existence of a 'realistic' interpretation, and that, Kripke claims, essentially, that his interpretation is to be construed as such, much more is needed by way of argument than the assertion that something is wrong with Kripke's notion of facts.

Dummett makes one further point:

> Kripke does not draw his distinction between the a priori and the necessary in the way that the distinction between epistemic and ontic necessity has here been drawn. On the account sketched above, epistemic necessity is a stronger notion than ontic necessity: a statement may be ontically but not epistemically necessary, but the converse could not occur. Kripke, however, claims the properties of being a priori and being necessary to be quite independent: not only may a statement be necessary though not a priori; it can also be a priori without being necessary.

But all this shows is that Kripke's notion of metaphysical necessity is not the same as Dummett's ontic. It gives no reason to suppose that Kripke's interpretation is not an equally valid "realistic" one. Indeed, we even saw that a slight variation of Dummett's ontic notion yielded a priori contingent statements.

The arguments considered so far in this section are not, cumulatively even, decisive against Dummett. They do cumulatively show, however, that if we must come down on one side in the dispute, it is Kripke's.

IV

Implicit in Dummett's discussion is an argument with, apparently, considerable force. Granting, for argument's sake, that Kripke's metaphysical/epistemic distinction can be drawn, Dummett might yet argue that it is only the epistemic context which is relevant to a term's meaning. Thus Dummett says:

> Even if this distinction were the right one to draw, it is plain that it is the notion of epistemic possibility that is required if we want to represent sense as a function from possible states of affairs to reference. Sense is (to repeat again) a cognitive notion.....(134).

Once this is granted, it is a short step to the claim that proper names can be synonymous with the definite descriptions which fix their reference. Within epistemic modal contexts, Kripke's arguments are disarmed for the behavior of names and descriptions are not distinguishable. Thus (a few lines later and concluding his argument):

> . . . then the way is open to consider even a proper name as a flexible designator: that is, to consider what object, if any, would, in a given possible world constitute its referent, if that referent were determined in the same way as is done in the real world ('Possible' here must mean 'epistemically possible': that is the only relevant notion when we are concerned with the epistemic question what we grasp in grasping the use of a word (pp. 134-135).[16]

---

[16] The first Dummett quote immediately above represents the idea that even if there is a distinction between metaphysical and epistemic necessity (which is not a distinction of scopes), it is only epistemic (de dicto) necessity which counts for a term's meaning.

And indeed peppered throughout Dummett's argument are suggestions similar to the above.

Before proceeding, let me make explicit a point contained in the above; and already mentioned by me somewhat obliquely. For both Dummett and Kripke, the truth conditions for an epistemic assertion of necessity are determined by what the terms in the sentence would refer to if used in an epistemically similar situation; or equivalently, by determining what terms would refer to if used with the same reference-fixing device. From such a perspective, 'Moses' might refer to someone other than what it does in the actual world (which Kripke would grant, given that we're talking about epistemic necessity).

Now Dummett takes the epistemic/metaphysical distinction to be really a de dicto/ de re distinction; epistemic is equivalent to de dicto. But taking a term nonrigid is also taking it within the scope of a modal operator. So saying a term is nonrigid (assume it is the only referring term in the sentence) is equivalent, for Dummett, to saying the assertion of modality, of which it is a part, is read

---

(Fn. 16 continued)

By the time Dummett gets to this second quote, he has once more reiterated that Kripke's distinction is really one of scopes. His point seems to be that in either case it is only the epistemic de dicto context which is relevant to a term's meaning.

epistemically. And this entails that taking a term as nonrigid means that its reference in any possible world is what the term would refer to if it were used in that world with the same reference-fixing device.

And this is precisely what Dummett indicates. Thus, in the prior long quote regarding rigid and nonrigid designators, he said:

> For definite descriptions, there is no distinction between their meaning and the way their reference is determined; in other words, the way in which the reference of a definite description is determined in the real world is carried over into each particular possible world.

And slightly later, "To assign a term a reference, varying from one possible world to another, is just to take it as having, in each possible world, the reference which it would bear in that world" (emphasis mine); and later on in the quote, "Thus, when Kripke says that it could not be true that: the teacher of Alexander didn't teach Alexander, he is intending to convey that, within any possible world, it would never be true to say, 'The teacher of Alexander didn't teach Alexander.'" And finally, our previous quote, " . . . then the way is open to consider even a proper name as a flexible designator; that is, to consider what object, if any, would, in a given possible world, constitute its referent, if that referent were determined in the same way as is done in the real world."

Having made this point, let us return to the original question of whether the epistemic context is the only rele-

vant one for a term's meaning.

To appreciate the strength of this contention, it is instructive to examine certain views of Putnam's; views which Putnam takes to be essentially equivalent to Kripke's:

> Let W_1 and W_2 be two possible worlds in which I exist and in which this glass exists and in which I am giving a meaning explanation by pointing to this glass and saying 'this is water'. (We do not assume that the liquid in the glass is the same in both worlds.) Let us suppose that in W_1 the glass is full of H_2O and in W_2 the glass is full of XYZ. We shall also suppose that W_1 is the actual world and that XYZ is the stuff typically called 'water' in the world W_2. (So that the relation between English speaker in W_1 and English speaker in W_2 is exactly the same as the relation between English speakers on earth and English speakers on Twin Earth.) Then there are two theories one might have concerning the meaning of 'water'.
> (1) One might hold that 'water' was world relative but constant in meaning (i.e., the word has a constant relative meaning.) In this theory 'water' means the same in W_1 and W_2; it's just that water is H_2O in W_1 and water is XYZ in W_2.
> (2) One might hold that water is H_2O in all worlds (the stuff called 'water' in W_2 isn't water) but 'water' doesn't have the same meaning in W_1 and W_2.
>
> If what was said before about the Twin Earth case was correct, then (2) is clearly the correct theory. When I say 'this (liquid) is water', the 'this' is, so to speak, a de re 'this'--i.e., the force of my explanation is that 'water' is whatever bears a certain equivalence relation (the relation we called 'same_L' above) to the piece of liquid referred to as 'this' in the actual world.
>
> We might symbolize the difference between the two theories as a 'scope' difference in the following way. In theory (1), the following is true:
>
> (1') (For every world W)(For every x in W), (x is water ≡ x bears same_L to the entity referred to as 'this' in W.)
>
> (2') (For every world W)(For every x in W)(x is water ≡ x bears same_L to the entity referred to as 'this' in the actual world W_1):

> (I call this a 'scope' difference because in (2') the entity referred to as ' "this" ' means 'the entity referred to as "this" in the actual world', and has thus a reference independent of the bound variable 'W'.)
>
> Kripke calls a designator 'rigid' (in a given sentence) if (in that sentence) it refers to the same individual in every possible world in which the designator designates. If we extend the notion of rigidity to substance names, then we may express Kripke's theory and mine by saying that the term 'water' is rigid (pp. 230-231).

It is interesting how close Putnam's analysis comes to Dummett's; even to the extent of adverting to scope distinctions. What divides them is that Putnam opts for (2) as the correct analysis (at least for terms like 'water') and contends thereby that difference in extension is 'ipso facto a difference in meaning' and gives up 'the contention that meanings are concepts or indeed mental entities of any kind'; while Dummett, naturally, will hold precisely the reverse; he will opt for (1) claiming that meaning is a cognitive notion (and that a difference in extension is not ipso facto a difference in meaning.[17]

Dummett would, we can imagine, argue as follows: Putnam's analysis confirms my claims; the rigid/nonrigid and epistemic/metaphysical distinctions are distinctions of scope. And even according to Putnam, if meaning is taken as a cognitive notion, (1) above is the correct analysis; and the only sense of meaning I understand (and certainly

---

[17] When I talk of (1) and (2), I should be construed as talking of the schemata of which (1) and (2) and (1') and (2') are instances.

Frege's meaning!) is the cognitive one. Further, it should be clear that my opting for (1) is equivalent to my claim that the only modal context relevant to a term's meaning is the epistemic.[18] Hence, there is, once again, no basis for distinguishing names from descriptions.

There are here two questions which need resolving. The first is whether or not (1) as opposed to (2) is the correct meaning analysis for referring and natural kind terms (including rigid designators). This question by its very nature is not amenable to definite resolution, but evidence can, I believe, be brought for coming down on one side.

The second question is whether our decision to analyze according to (1) or (2) (1' or 2') is equivalent to the decision of whether to take a term (e.g., 'water') as rigid or nonrigid (and equivalent also to the decision of whether to take de dicto modal assertions in which the terms occur as epistemic or metaphysical). The validity of this equivalence is crucial for the argument we are expounding on Dummett's behalf. For that argument is that if we take a term within an epistemic context (i.e., if we analyze according to (1) or (1') (see Fn. 18), we are thereby treating

---

[18] To see the equivalence of (1) and (1') with Dummett's analysis of the epistemic context, replace in (1') 'water' with 'Jack the Ripper' and ' . . . bears the same$_L$ as the entity referred to by "this" ' with 'is the famous London murderer'. We might also point out that Putnam's 'constant relative meaning' comes to the same as Dummett's 'the way the reference is determined in the actual world is carried over into each world' (from the long quote in Section II).

it nonrigidly; consequently, if we restrict outselves to epistemic contexts, there is no distinction between names and descriptions--they both are used nonrigidly. Now if we show that giving (1) as the correct analysis of a referring term does not entail that it is being used nonrigidly, then this argument, persuasive though it appears, will fall.

As to the first question, let us consider the three terms: 'I', 'Water', and 'Jack the Ripper', (or any other proper name). These three are arranged in ascending order of plausibility of having (2) as the correct meaning analysis.

Not for a term such as 'I', it is clear, as Putnam admits, that the correct analysis is (1). Merely because in some other world 'I' can be used to refer to a person other than that which it does in the actual world is not in itself reason to contend that 'I' is being used with different meaning in the two possible worlds; just as similar remarks hold of two uses of 'I' in the actual world.

In the case of a natural kind term such as 'water', there is more room for dispute. My own intuition coincides with Putnam's and independent reasons can, I believe, be adduced. But the issue is not clear and (1) might without much straining be plausibly defended by some.

In the case of proper names, however, it seems clear that contra Dummett, (2) must be the correct analysis. That is, if in some other possible, or actual, world, we use a

name orthographically and phonetically the same as that which we use in the actual world, with the same reference-fixing device, it would still standardly mean something other than what it actually does if it were used in that world to pick out some other person (that is, if the idea of a name having meaning makes sense at all--which Dummett thinks it does).

This point is made more persuasive if one considers the actual use of a name, 'Sam', in two different contexts (on distant planets, say). My utterance and an Alpha Centurion's of 'Sam is wise' have different meaning--after all, when queried I must respond that I really don't know what he means; who is this Sam he is referring to? Thus, the two uses of 'Sam' (even using the same reference-fixing device) have different meaning (recall we are assuming that it makes sense to talk of names having meaning).

The above argument (if it can be called an argument at all) is hardly conclusive. What it shows is that on the face of it, (1) is not the correct meaning analysis of proper names. But one already committed to Dummett's view will hardly be moved by its force.[19]

My response to the second question will, I believe, be less easy to dismiss. On its basis, we can see that even if

---

[19]Or, it might be responded, since (2) is equivalent to (2'), the best the argument can show is that according to "ad hoc" conventions of scope, proper names go out of the scope of the modal operator. I do not find this line of argument persuasive. In any case, we shall see below that the equivalence of (1) and (2) with (1') and (2') with rigidity and nonrigidity can not be supported.

it were conceded that for natural kind terms and proper names (1) is the correct analysis, Dummett's argument will not stand.

We have seen that for Dummett to say that a term is used nonrigidly is to say that the same "cognitive meaning", i.e., the same reference-fixing device, is carried over to the term's use in other possible worlds--that is (1) and (1') (Dummett's epistemic context) are the correct analyses. This is a crucial premise for his argument for the next step is that even proper names should or can be analyzed according to (1) and (1') and hence, can be construed as nonrigid. Now all that now needs to be produced to show the premise under consideration is a term which is clearly rigid and clearly has (1) or (1') as the correct analysis. Such a term is 'I'.

As Putnam rightly conceded, pronouns such as 'I' are best analyzed by (1) and (1'). But as Putnam also rightly argues, 'I' is rigid (or an 'indexical' as he calls it); when used modally it refers to the same object in every possible world--as a moment's thought will show.[20]

---

[20] Consider 'It is possible that I (you) was (were) the 37th president'.
Also in Fn. 12 of his preface, Kripke says:

> For example, some philosophers would assimilate proper names to demonstratives. Their reference varies from utterance to utterance the way that of a demonstrative does. This does not affect the issues discussed, since the reference of a demonstrative must be given for a definite proposition to be expressed. Although I did not discuss the question in the present monograph, of course it was part of my view (p. 49, n. 16) that 'this', 'I', 'you', etc., are all rigid (even though their references

In Putnam's terms, this means that either (1) and (2) are not always equivalent to (1') and (2'), or that (1') and (2') are not definitive of rigidity and nonrigidity. The crucial point for Dummett is that even if it is granted that (1), and what Dummett and Putnam take to be equivalent (1'), is the correct meaning analysis of referring terms--which is far from obvious in the case of natural kind terms and proper names--the distinction between terms used rigidly and nonrigidly is not reducible to the distinction between (1) and (2).

This point is brought out most forcefully by our example of 'I' (or 'you' etc.). We can consider what the term would refer to in some other possible situation if used with the same meaning as it has in the actual world; and it certainly can be so used to pick out someone other than it actually picks out. But this clearly does not show 'I' is nonrigid the way, say 'the first president' is nonrigid. Thus, we see that even if Dummett is right in contending that (1) is the correct meaning analysis for all referring terms, that is not to say that the term is necessarily being used nonrigidly. So the argument we are ascribing to Dummett in its present state will not stand.

(Fn. 20 continued)

obviously vary with the context of utterance). The rigidity of demonstratives has been stressed by David Kaplan.

The overall conclusion of my discussion of Dummett's critique, is that it does not stand. Where the issues have been joined, Kripke has come out on the winning side.

CHOMSKY'S CRITIQUE

Chomsky's Critique

I wish next to discuss some recent criticisms, due to Chomsky, of Kripke's account of essential properties.[1] Chomsky's claim, briefly put, is that what an object's essential properties are is relative to our categorizations of that object, and that Kripke's purported examples of de re attribution of essential properties are actually cases of de dicto necessity.[2]

This claim has significance for Kripke's mind/body argument (though Chomsky himself does not take up the point). If, as Chomsky claims, things have essential properties only relative to our categorization of them, then it would be open to us to assert that it is essential to particular pains that they are pains only insofar as they are categorized as such. But we can, and do, equally well categorize them in neurophysiological terms; and as such, they are not essentially mental entities at all. This, it appears, would be enough to vitiate Kripke's argument.

I. Chomsky's argument proceeds by first claiming that giving an object a name (and that object having that name) presupposes a set of categorizations and conditions.

---

[1] In, Reflections on Language, Pp. 44-52. All references to Chomsky will be to this.

[2] There have been many arguments in the literature along similar lines. I discuss Chomsky's because, among other reasons, of its intrinsic interest and the caution of its claims.

First, to attach a name to an entity, that entity must be "naturally nameable"; this involves satisfying conditions having to do with spatiotemporal contiguity, Gestalt qualities, functions within the space of human actions, and so on (e.g., any mundane object such as a tree is "naturally nameable"). Furthermore, "in determining that an entity is a nameable thing, we assign it to a 'natural kind' that might be designated by a common noun, a 'sortal predicate' . . . . This assignment involves assumptions about the nature of things named--some conceptual and some factual."

Besides these conceptual and factual assumptions drawn from "the structure of common sense understanding," there are further conditions imposed by the "cognitive structure of language." Names are either personal names, place names, color names, etc., each with its own intended domain of discourse. There are no "pure" names.

Though I sympathize with much of this, it seems, as it stands, clearly false. Indeed, any constraints on what can be named, or the specific beliefs and categories that naming presupposes, is, I believe, bound to be wrong. We can, in theory, name anything and everything (though perhaps not all at once); there are no restrictions on what can qualify as a "nameable" thing. I now name the entity which consists of the French Revolution, the first twenty years of Chomsky's life, and the far side of the moon:

"Th 24." Is there anything to prevent me from doing so? And what natural kind does this entity belong to?

There is no point, to be sure, to naming such an entity (as there generally would not be for such odd entities); such things tend to be intrinsically uninteresting and names for them would tend not to gain currency. But there is no reason why they could not gain currency, and certainly, there is nothing which prohibits their use as names.

This point is, perhaps, circumventable. We might, for instance, contend that such recherché objects are "derivative," in some sense, on more "basic" objects (dealing somehow with Goodmanian problems), and these more basic objects are "naturally nameable." Further, in the case of most names, there are indeed factual and conceptual assumptions attendant upon those terms' use.

But even so, it seems clear that there need not, in every case of naming, be some particular conceptual and factual assumptions that we make about the object named--some particular category (or natural kind) which we must place it in. The strongest assertion we can plausibly make, and the element of truth in Chomsky's idea, is, I think, that in standard cases of giving something a name, and in using names generally, it is presupposed that the object named belongs to some kind or other. That, however, is a much weaker claim than Chomsky's.

Thus, to give an example, imagine a quiz show in

which I am the master of ceremonies and announce that there is some object or event--unknown even to myself--which has either occurred in, or been placed in, the adjoining room. This entity I call 'Arthur'. The task of the contestants, I declare, is to determine the identity of Arthur. Questions may be of the nature of:

(1) Is Arthur an event or object?

(2) Is Arthur animal, vegetable, or mineral?

'Arthur' strikes me as a perfectly understandable and natural name, though neither I, nor the contestants, have any idea what kind of thing Arthur is.

This is an example of an initial naming situation. In the majority of instances, however, we use names which are already in the linguistic community's lexicon. And in such a situation, my point will be even stronger; for in such a case, it appears even clearer that our use of a name is not dependent on our having a particular set of beliefs about the nature of the object named.

Kripke offers the example of a mathematician's spouse who does not know whether the mathematician's use of 'Nancy' was to refer to a woman or to a Lie group. More mundane examples are not hard to come by. At a cocktail party, I discern in my friend's chatter constant reference to (Michaelangelo and) Sagitareus; in complete ignorance of the topic of conversation, but wishing to seem erudite, I say--to the consternation of my friends--"Sagitareus is huge." Or, in a similar situation, I say, "The Holy See

stinks"; not knowing whether I am referring to a person, a lake, or whatever.

In these cases, I intend (successfully) to refer by the use of the name to whatever the others in my small group referred to by the use of the name. Kripke's analysis of reference seems correct in its broad outlines. In most cases of using a name to refer, one must intend that, by one's use of the name, one refers to whomever, or whatever, was referred to by either the person(s) whom one picked up the name from, or society at large, or an appropriate set of specialists, etc. (unpacking the 'etc.' is the hard part). There need not be any <u>particular</u> categorization on the user's part--though there usually is--in either initial namings, or in ordinary situations.

II. Chomsky uses his conception of names just described to claim that Kripke's notion of de re necessity--that is, the attribution of essential properties--is just a case of de dicto necessity in disguise. Thus, Chomsky considers the two sentences:

(1) Nixon won the 1968 election.

(2) Nixon is an animate object.

(2) would, in Kripke's view, be an attribution of a necessary property to Nixon; assuming Nixon is, in fact, animate, being animate is one of his necessary properties.

This Kripkean intuition of necessity is due, claims Chomsky, to the fact that statement (2) is (approximately)

synonymous with:

(3) The person Nixon is animate

since 'Nixon' is a personal name. Thus, (2), as with all other putative cases of attributions of necessary properties, is, as is (3), a case of de dicto necessity.[3]

> This necessary truth (i.e., (2) and (3)) may be grounded in a necessary connection between categories of common sense understanding, or an analytic connection between the linguistic terms 'person' and 'animate'. Under any of these assumptions, we need not suppose that an essential property is assigned to an individual, Nixon, apart from the way he is named or the category of common sense understanding which he is assigned.[4]

Two points: First, Chomsky's argument depends on the claim that (2) and (3) are synonymous. That depends on the claim that 'Nixon' is synonymous with 'The person Nixon'; and that claim, I take it, entails that if we did not use 'Nixon' (and similarly with other proper names) with the sense of 'the person Nixon', if, say, we had suspended judgment on what kind of thing 'Nixon' referred to, then we would be using the name with other than its ordinary sense (or ordinary usage, or whatever; those who do not believe names have sense should rephrase the above point appropriately).

---

[3]Thus, Chomsky's argument is the reverse of Dummett's. Dummett argues that apparent claims of de dicto metaphysical necessity are really de re attributions of necessary properties, while Chomsky argues that apparent de re attributions of metaphysically necessary properties are really de dicto attributions of necessity.

[4]Chomsky, Pg. 46.

The brunt of my discussion in section I was just that Chomsky is wrong in this claim. Names can be used in a perfectly ordinary manner (without changing their meanings or usage), without there being any assumptions about the kind of thing named. So, if my argument in section I was correct, Chomsky's argument will not work here.

To this first point, Chomsky might respond that he can grant that there may be uses of proper names which do not presuppose a particular set of beliefs; nor that the object named belongs to a certain category. Nevertheless, in this particular case, we find an aura of necessity surrounding (2): Nixon is animate because, in point-of-fact, we standardly do assume that 'Nixon' is a personal name; consequently, (2) is more or less synonymous with (3).

If we did, Chomsky may continue, attempt the psychologically difficult feat of using 'Nixon' as a "pure" name, then (2) would no longer seem to us to be the attribution of a necessary property. So, the necessity we discern in (2) is due to its (de facto) equivalence to (3), and hence, its necessity is de dicto.

And Chomsky does propose an argument that more or less comes to this.[5] Thus, he considers the possibility that our

---

[5]The difference is that Chomsky does not grant that there can be "pure" names in a natural language--he merely considers the possibility for the sake of argument.

As for the possibility of "pure" names in natural languages, Ned Block has suggested to me that one can imagine someone writing a story that begins, "As K awoke one morning from uneasy dreams, he found himself transformed in his bed into a gigantic insect." Notice, too, the possibility of conceiving of such a situation might be used in an argument about what properties are essential to individuals; that one can, therefore, make essential attributions using "pure" names (See below).

name for Nixon did not, in fact, mean 'the person Nixon'--that it is a pure name, 'N' say. In that case, the analog of (2), (2') N is animate, would not have the aura of necessity (2) has; for (2) has it only because 'Nixon' is a personal name, which 'N' is not. So, as above, the necessity of (2) is de dicto. If we do not read (2) equivalent to (3), but read it equivalent to (2'), its necessity will not be perceived.

So far, there is nothing Chomsky has said that would have any force for one who is already committed to essential attribution. To such an individual, there is still nothing problematic with the contention that 'N is animate' attributes an essential property--no more so (or less so) than in the case of 'Nixon is animate'.

What Chomsky says in the way of support for his contention is,

> Within this new invented system (i.e., of pure names), divorced from language and common sense understanding, we have no relevant intuitions to guide us, I believe. Thus, we might want to say that there is no way for the thing, N, to have been anything other than what it is. For then it would have been a different thing; thus, in an uninteresting sense, all its properties are 'essential'. Or, we might want to say that any of its properties might have been other than what they are. The expository recourse to distinctions between (1) and (2) (for example), between what might have been and what could not have been otherwise, is no longer available to us within the new system we are now imagining.

This argument is not persuasive. Certainly, the two essentialist positions described are attributable whether we use "pure" names, as above, or whether we use personal

names, such as 'Nixon'. In either case, we can contend that all or none of an object's properties are essential. Indeed, such views have, in fact, been forwarded by various philosophers (e.g., Leibnitz the former; Russell the latter). I see no distinction of significance that can be drawn between "pure" names and ordinary names on this score.

There is, to be sure, <u>some</u> difference between 'N' and 'Nixon'. If we did not know that 'N' standardly referred to a person, we would not know (2') to be even true--a fortiori that it truly attributed a necessary property. But even if we did not <u>know</u> this to be so, it would, nevertheless, <u>an essentialist would claim</u>, still be so. And given that we <u>do</u> know that 'N' refers to a person--no matter how that information was acquired--we would thereby know that an essential property had been truly attributed.

So the difference between 'N' and 'Nixon' is of no significance as far as essential attribution goes; and we can not conclude that (2) is necessary de dicto. To conclude there is a relevant distinction between 'N' and 'Nixon', one must assume that the necessity of (2) is de dicto--an assumption the essentialist is obviously not prepared to make. (I should point out that I am not here purporting to offer positive arguments for essentialism; merely contending that Chomsky's arguments do not affect one already committed to it.)

One last consideration is this: need I here defend what I have been assuming without argument, that being animate is a necessary property of Nixon? There are those who would deny this (though I supposed that anyone who grants that there are nontrivial essential properties will grant that this is one of them). This undefended assumption is, in the present context, legitimate because Chomsky is not here taking issue with the particular properties Kripke claims are essential; he is contending that the notion of essential property--independent of how the object is categorized or referred to--is incoherent; that the modal status of (2) depends on its containing 'Nixon' and not 'N'.

My point has been that Chomsky's example does not show this at all (unless, of course, one already assumes that essential attributions are incoherent); that one can ascribe necessity just as easily with "pure" names as with ordinary names. He has not shown that there is any relevant difference between these sorts of names; nor that there is anything incoherent about essential properties as interpreted by Kripke.

I now turn to the second of my two points. Kripke might certainly admit without difficulty that in many cases, what partially fixes the reference of proper names are descriptions such as: 'is animate', 'is a person', and so on. This, in Kripke's view, would entail that, e.g., 'Nixon is animate', can be known a priori; for we would

then take 'Nixon' to refer to what (and only what) is an animate thing. That this might be the case does not, as Kripke points out, affect his contentions about necessary properties.[6]

But Chomsky is not here purporting to point out the possibility of such a priori knowledge. He quite rightly claims that 'Nixon is animate' can not be known a priori. (Kripke would agree and say that this shows that 'is animate' is not sufficient to "fix the reference" of 'Nixon'.) But if, as Chomsky grants, it is not a priori, how can it be synonymous with, 'The person Nixon is animate', which apparently is--especially if the necessity that Chomsky discerns in the latter statement is grounded in the relation between 'person' and 'animate'.

It strikes me that I am missing something here, but I am not sure what. Chomsky says later, " . . . suppose that we were to discover that the entity named 'Nixon' was, in fact, an automaton so that (2) was false. We might then conclude that the personal name 'Nixon' had been misused (so we now discover) . . . ." It is true that if 'Nixon' meant 'the person Nixon', then 'Nixon' would have been misused. For 'Nixon' (synonymous with 'the person Nixon') would purport to pick out something which is a person where there isn't any such.

'Nixon', however, does not seem to be such that it would be misused if there turned out to be no person who

---

[6]See NN, Pp. 351-352, Fn. 58 for instance.

was Nixon; we would just say (at least, so it seems to me), that Nixon is not animate (we have no found out).

Further, the lack of a person Nixon does not affect the epistemic status of the assertion 'The person Nixon is a person' (just as the a priori status of 'The man living upstairs lives upstairs' is not affected by the possibility that there is no man upstairs); but the epistemic status of 'Nixon is a person' is affected by the possibility of the lack of a person Nixon--it is this possibility which makes the statement a posteori.

In the same vein, we might further ask of the statement, 'Nixon is not a person', what its epistemic status is. This, in Chomsky's view, should be synonymous with 'The person Nixon is not a person', which is unequivocally, logically false. Yet, there is a clear sense in which 'Nixon is not a person' is not logically false. So how are the two synonymous?[7]

These two points show, I believe, that Chomsky has shown nothing incoherent with Kripke's views; and that his inferences are open to serious question.[8]

---

[7]Indeed, what does Chomsky do with 'The red table might not have been red'? This clearly has a true reading--but no true de dicto reading. So it must have a true de re reading.
Certain philosophers with views similar to Chomsky's (e.g., Plantigina and Kaplan), attempt to deal with this sort of problem. Their views are refinements of the contention that sentences as the above are to be analyzed as "There is a term α such that α refers to the red table and 'α is red' is possibly false."
Chomsky's analysis is different, however. The assertion that Nixon is necessarily animate is analyzed in terms of the synonymy 'Nixon is animate' and 'The person Nixon is animate'.

[8]Kripke need not deny any of Chomsky's insights about common sense and conceptual categorizations. One can easily imagine him granting that, though it is not the last word, one's common sense and conceptual categorizations have a lot to do with what kinds of things we think there are, and consequently, with what essential properties we think there are.

III. Chomsky makes his above points again while discussing some of Kripke's examples of essential properties, i.e., having a particular set of parents and being a table. In the latter case, he presents two further arguments.

What looks like a perfectly normal table might have been, contends Chomsky, designed to be a hard bed and actually used as such.

> Surely, we would then say that thing is not a table but a hard bed that looks like a table. But the thing is what it is. Neither a gleam in the eye of the inventor nor general custom can determine its essential properties, though intention and function are relevant to determining what we take an artifact to be.

Again, the point is that what essential properties an object has depends on how we categorize it and is not intrinsic to the object itself. But, as Chomsky himself says, what we have here is merely a case of a bed which looks like a table. If intention and function are relevant to determining what we take an object to be, why balk at the contention (which I think true) that what an artifact <u>is</u>, is partially determined by the producer's intention and by the object's function. These may, in some sense, not be any of the object's physical properties, but so what? As Chomsky notices, such things as constitution, structure, and agent responsible for generation, are among Aristotle's and Kripke's list of kinds of essential properties. Function and intention, when it comes to artifacts, should likewise be included. After all, what an artifact <u>is</u>, as the

table example amply shows, depends on its function and intention. Nothing is a carburetor unless it can function like a carburetor (when not broken), to take a venerable example; it is not just that it wouldn't be described as a carburetor--it really wouldn't be a carburetor.

Of course, to the extent that what an artifact is depends on what it was intended to be, that artifact's essential properties depend on human cognition, broadly construed. This fact hardly makes Chomsky's point; all it shows is that sometimes what an object is--what its essential properties are--depends on human intentions. It does not show that all of any object's essential properties are relative to our categorization of that object and that it is incoherent to think otherwise.

Chomsky's other move is to consider possible alien creatures who say that a table "would have been a different thing had it been nailed to the floor, though it could have been other than a table." To them, as opposed to us, immovability would be an essential property of objects while being a table would not. And there appears to be no way of adjudicating between the two ways of looking at the world. Once again, it seems that what essential properties an object has depends on how one categorizes it.

How one would indeed adjudicate between such conflicting ontologies is a difficult and tantalizing question. One answer is Chomsky's: that there is no genuine question as to who is right; there are just different ways of categorizing.

But this is not the only possible answer. The best solution, to my eyes, is the liberal one of admitting both ontologies, and both sets of essential properties--when, that is, there is no further objective way of deciding between the two. In the present case, this entails admitting earth tables (and other artifacts) which may or may not be movable objects, and alien movable (and immovable) objects, which may or may not be tablelike. Each set of objects will have its own set of essential properties.

We can even allow that one hunk of matter might be the matter of an instance of both these kinds of things; we would have two things in the same place, at the same time; however, Locke's proviso: of the same kind, would not be satisfied.

Or perhaps, this liberal move is to be eschewed.[9] A different view, distinct from Chomsky's, might be just that in fact, we and the aliens would disagree about what kinds of things there are and what essential properties these have. This, in itself, does not entail that there are not (description and categorization independent) kinds of things or essential properties. For what is this disagreement about, other than what kinds of things there are (and what essential properties there are)? If there are no categorizationally independent kinds of things and essential properties, where then is the disagreement? All hands can certainly agree

[9]Perhaps we could imagine some situation in which the ontologies were truly incompatible.

that each race believes, on the basis of its categorizations, different kinds of things to exist than the other--but that is not the issue. The question in this analysis is, who, if anyone, is right? That is, what kinds of things and essential properties are there?

Again, my point here has not been to offer positive arguments for essentialism; rather, I wished to show that Chomsky's examples need not perturb the essentialist, for the essentialist can analyze these examples in quite reasonable ways, compatible with his views.

IV. Chomsky claims that his views do not commit him to any particular metaphysical outlook. He is thus able to hold aloof from most of the metaphysical claims and counterclaims surrounding the notion of essential attribution. Perhaps it is for this reason that his arguments do not, as we have seen, have any force for those already committed to essentialism.

V. If all this is correct, then the possible argument against Kripke's mind/body argument with which we started, will not work. There is still, however, an analogous argument, in the material mode, as it were, that might be made. I shall take this up in the next section.

---

[10]Essential properties need not, of course, enter into this disagreement at all. The conflict can be described as one of varying ontologies, without any mention of essential properties. But it was Chomsky's point that in the situation described, we apparently have different essential properties and that, therefore, the notion of essential property independent of categorization, is incoherent; my point has been that it does not follow.

FELDMAN'S CRITIQUE

Feldman's Critique of Kripke

This section will discuss Feldman's response to Kripke's antimaterialist argument; a response which has received wide currency and is often taken to be conclusive.

Feldman is concerned with Kripke's denial that, "each person is identical with his or her body," and that, "each particular mental event or state is identical to some corresponding physical event or state."[1] His aim is to show that (1) Kripke's arguments are open to serious objections; and (2) Kripke's principle that true identity statements containing rigid designators are necessarily true, is not, pace Kripke, relevant to arguments concerning mind/body identity. My concerns shall be mainly with the first of these aims.

There are three basic strategies in Feldman's attack. In the case of persons and bodies, they are the claims that:

(1) Persons are not essentially persons.
(2) Bodies are not essentially bodies.
(3) The contingency felt to obtain between persons and their bodies can be analyzed in a manner (given below) innocuous to materialism.

---

[1] See Feldman, "Kripke on the Identity Theory," Journal of Philosophy Vl, LXXI, No. 18 (Oct. 24, 1974): pp. 665-676; and, "Kripke's Argument Against Materialism," Philosophical Studies XXIV, 6 (November 1973): pp. 416-419.

When applied to the mental and physical states (events) case, these become:

(1') Mental states are not essentially mental states.
(2') Brain states are not essentially brain states.
(3') The contingency felt to obtain between mental states and brain states can be analyzed in a manner innocuous to materialism.

In the person/body case, Feldman concentrates on (3), though he does have some discussion of (1) and (2).[2] In the mental/brain event case, Feldman concentrates exclusively on (1'). He mentions in passing only (2'); and does not even

---

[2](1) is used primarily against Kripke's nonmodal (Feldman's term) arguments. Kripke mentions, though does not unconditionally endorse, the argument that since (dead) bodies exist without the person to which they are putatively identical existing, the bodies are not identical with those persons.

Feldman contends that the materialist would respond that persons--that is, certain bodies--are not essentially alive, or not essentially persons. A person can exist without being alive or a person, just as a particular red rose might exist without being red. (Though the argument mentioned by Kripke is itself nonmodal, the objection introduces modal notions.)

A materialist could extricate himself as Feldman suggests, but if being a person is not an essential property of persons, then what is? This point does not depend on the coherency of essential attribution. It can be simply put by saying that people, and not their bodies, go out of existence when they die. Elvis Presley is no more, though many still make pilgrimages to his body.

There is an adequate reply to this sort of nonmodal argument, one mentioned by Kripke himself. Perhaps Feldman makes the response he does simply because it parallels his analyses of Kripke's other arguments.

The reply Kripke suggests is that analogously to the case of the statue and the matter of which it is made, or the ship and the planks of wood of which it is composed, a person is a body with a certain physical organization--one that it does not have when it is dead.

There is a sense in which the ship is not identical with the planks of wood; the planks can exist when the ship does not; but there is equally a sense in which the ship is the planks ("nothing over and above" the planks, as Kripke says). The added element of organization is not some further ob-

do so much for (3'). Why Feldman chose to concentrate in particular on (3) and (1'), I do not know.

Let me note that my discussion of (1') will be relevant to Lewis' materialist claims. Our previous examination of Lewis focused on his type/type views; Kripke's arguments, however, will, if valid, apply also to Lewis, even construed as espousing a token/token view.

The issue between Kripke and Lewis appears to reduce to the question of whether we can take 'pain' to be a nonrigid designator (See the discussion in Fn. 7, section 2 in the Lewis half of the thesis; see also Lewis' "Mad Pain and Martian Pain," in particular, Fn. 2). Or equivalently--Feldman, and probably Lewis, would assert--whether pain is essentially painful. But this is the claim of (1'), which we shall be discussing.

Lewis has, of course, independent reasons for adhering to (1'), that Feldman does not adduce; i.e., his entire framework and justification of theoretical and psychological definition. Of course, to the extent that we have previously questioned the correctness of these definitions, we also undercut Lewis' defenses here. Let me now turn to

---

(Fn. 2 continued)

ject that is added to the planks of wood--or the body of the person or the clay of the statue; it is their form. I think that a materialist would be perfectly happy with the contention that the relation of bodies to personsis the same as the relation of planks of wood to the ships whose material it is (whatever that relation precisely is). This, I think, is an adequate response to the 'nonmodal' arguments.

the particulars of Feldman's arguments.

The Person/Body Argument

Kripke's argument against person/body identity is (as we know) in outline: Let 'd' be a rigid designator of Descartes, and 'b' a rigid designator of Descartes' body. If b=d, then (since 'b' and 'd' are rigid) necessarily, b=d. But it is possible that b≠ d; therefore, b≠d. Feldman formalizes the argument as follows (the numbering is his):

(13) 'd' is rigid and 'b' is rigid.

(14) '◇(∃t)(Edt&¬Ebt)' is true.

(15) If α is rigid and β is rigid, and ⌜◇(∃t)(Eαt& Eβt) is true, then ⌜◇($\alpha \neq \beta$)⌝ is true.

(16) If α is rigid, and β is rigid, and ⌜◇($\alpha \neq \beta$)⌝ is true, then ⌜$\alpha \neq \beta$⌝ is true.

(17) 'd≠b' is true.

where 'Ext' means that x exists at time t.

The bone of contention, claims Feldman, is (14); or better, neglecting the time variable,[3] the claim that Descartes can possibly exist without Decartes' body. (From now on, I should be construed, when referring to (14), as referring to this claim.)

Kripke's claim is that one can not, as the above argument proves, accept (14) and then blithly proceed to

[3]The time variable just complicates matters and for the rest of the paper, I shall ignore it. (14) is then '◇(Ed&¬Eb)'.

assert the identity of Descartes with his body. But, responds Feldman, Kripke has merely asserted, not shown (14) to be true. And until he has done so, he has not proven the non-identity of d and b.

We may grant Feldman that Kripke has not proven his point. Much of his argument reads as if he is attempting to show that materialists can logically deny Kripke's conclusion; that it is open to them to deny, without inconsistency, one of the premises (and hence the conclusion). But that hardly constitutes a counterargument. Where, after all (outside of say logic or the philosophy of logic), do we have substantial philosophical theses which can be proven (formally)?

What Feldman must show, if his objection is to carry weight, is that it is reasonable for the materialist to deny (14). To merely assert that an intelligent materialist would notice that he must deny (14) is not the same as producing a reason for such a denial.

There is, after all, prima facie intuitive support for the contention that people might not be bodies. (Even such materialists as Armstrong concede the possibility of bodiless persons.) We all feel, it seems to me, the intuitive force of this contention; intuitions are, to be sure, often unreliable; but the burden of proof is surely on someone who denied what appears to be intuitively true: in this case, Feldman.

There is a further point to be considered. Even were Feldman's denial to (14) reasonable, still, Kripke's argument is, to a large extent, aimed at the "Contingent Identity Theorist" (henceforward, CITs) who do, it seems, accept (14), while at the same time, asserting the identity of d and b. Kripke's argument may then be looked at as a reductio ad absurdum of the CITs' position.

Feldman entertains this analysis of Kripke's argument, but rejects it on the grounds that the CITs do not, or should not accept (14).

As a matter of historical fact, this does not seem to be quite accurate. Exegesis, however, of the literature representing the CITs' view to determine what, if anything, was precisely meant, would clearly be to no purpose. If Feldman can come up with an adequate alternative to (14), which the CITs could have, or should have, espoused, and which does not, as does (14), entail that d≠b, then he will have an effective counterargument against Kripke. Kripke will no longer have a reductio ad absurdum against the CIT view--when properly formulated.

What Feldman claims the CIT should have asserted is:

(18) $(x)(t)(Pxt\supset(\exists y)(Byt\&Oyxt\&x=y)\ \&$
$\neg\Box(x)(t)(Pxt\supset(\exists y)(Byt\&Oyxt\&x=y)$.

where 'Pxt' means x is a person at t, and 'Bxt' means x is a body at t, and 'Oyxt' means x belongs to y at t.[4]

---

[4]Without the time variable, we have: $(x)(Px\supset Bx)\&\neg\Box(x)(Px\supset Bx)$; (19) below would become: $(x)(Px\supset(Ey)(By\&x=y\&\neg\Box x=y))$.

From (18), claims Feldman, one can not deduce (14); what would entail (14) is:

(19) $(x)(t)(Pxt\supset(\exists y)(Byt\&Oyxt\&x=y\&\neg\Box(x=y))$.

But (18) does not entail, asserts Feldman, (19)(nor (14), nor that $d\neq b$). (18) is true if everyone in the world is necessarily identical with his or her body, though in other possible worlds, there do exist pure spirits. That, however, would make (19) false. So, it appears that the CITs can, by accepting (18), assert, in a sense, the contingency of the identity of people and bodies; i.e., they may grant the possibility of a world with bodiless people, and still deny of any actual person that he might not have been a body. Thus, they may deny (14), and block Kripke's reductio.

Again, as a matter of historical fact, it strikes me that (19) is closer than (18) to what CITs actually said. To be sure, (19) is a silly thing to accept; but that is precisely the point of Kripke's attack.

We are, in any case, left with two pertinent questions. First, do the CITs, whether or not they so intended, have recourse to (18) and thereby escape Kripke's reductio? Second, is holding (18), whether or not CITs would or should do so, a reasonable response to Kripke's challenge? That is, might we not say that (18) is behind the intuitions of the contingent relation between persons and bodies, which I appealed to previously when throwing the burden of proof on

Feldman. If so, then appeals to such intuitions will be of no support to Kripke's claims; for even if valid, they will not entail the nonidentity of Descartes and Descartes' body.

Besides these two, I shall also discuss a third issue: Feldman's contention that the materialist could defend against Kripke by claiming that bodies are not essentially bodies; and a fourth: Feldman's contention that the notion of rigid designator is not relevant to questions of mind/body.

(1) In response to our first question, the answer must be no. If the CITs hold <u>anything</u>, it is that the identity of people and their bodies, and mental states with physiological states, is to be construed on analogy with the better known cases of "contingent" identity: e.g., that of water with H_2O, light with electromagnetic radiation, heat with molecular motion, or lightning with electrical discharge.

What I shall show is that in <u>these</u> cases, they would not, or should not, accept analogs of (18). Consequently, they would not, or should not, accept (18)--for again, the crux of their position is just that the analysis of person/body identity is to be construed analogously to that of other contingent identities.

Suppose some CIT were challenged by Kripke's arguments (and suppose he accepted, as does Feldman, Kripke's contention that identity statements containing rigid

designators are necessarily true if true at all); that is, suppose it were argued that if heat is identical with molecular motion, then it is necessarily so identical; that in a situation in which you do not have molecular motion, you ipso facto do not have heat, though you may have something which feels like heat. How would the CIT respond?

My conception of how he should respond is: You are correct, Kripke, in calling us to task for the formulation of our position. Yet our view is, we claim, basically sound; and you, yourself, have accurately formulated what we have wished to say: Given the way we pick out heat (that is, by the way it feels), 'heat' might have referred to something other than it does), something other than molecular motion. Our view of contingent identity is probably best put in terms of what you call "epistemic" modality: It is only contingently true that when we are in an epistemic state similar to that which we are in when we actually fix the reference of heat, that we refer to heat--i.e., molecular motion. This is what we mean when we say that heat is only contingently molecular motion. Analogous comments will apply to persons and bodies. We might be in the same epistemic state as we are actually in relative to Descartes and people in general, and yet there be something other than Descartes or people.[5]

---

[5]So it appears that the CITs have in any case a reply to Kripke. Kripke indicates that the sort of reply

There is no way of knowing that such would actually be the CITs' reply, but it is the most plausible line to take.[6]

But according to Feldman, this should not be the CITs' response. For, to respond thus, would not be to respond in a manner analogous to (18). An analog of (18) would be something like (18'): (x)(Hx↔Mmx(&¬□(x)(Hx↔Mmx); where

---

(Fn. 5 continued)

mentioned here would suffer from problems analogous to those suffered by a similar response in the pain state/brain state case, but he does not say precisely what that problem is in the present case.

Perhaps, he has in mind something like this: It seems true to say, "I might not have been a (or this) body." This intuition can not be explained away in the manner described above: We can not say we are imagining a case in which 'I' refers to some other entity--for ex hypothese it is I who am doing the referring, so it can only be me that I am referring to; consequently, it must really be I who am possibly not a body. Something like this is what I imagine Kripke had in mind, though I am not sure.

[6]Notice the following quote from Armstrong (The Monist: (April 1972) Vol. 56 #2, "Materialism, Properties, and Predicates"). In this quote, he distinguishes properties from predicates, the latter being linguistic entities.

> We can pick heat from other sensory qualities but that does not mean that perception (or introspection) yields any awareness of what the quality is. What it is, we first learn later as a result of the labor of scientists, who make a theoretical identification of heat with mean kinetic velocity of molecules . . . . We have seen that cases where two logically distinct predicates apply in virtue of the same property . . . . 'Anger is a certain sort of firing of neural circuits' involves two logically distinct predicates, both of which apply to the objects they apply to in virtue of the same property. But the two predicates do this in a semantically different way.

'H' is shorthand for 'is (an instance of) heat', and 'Mm' is shorthand for 'is (an instance of) molecular motion'. That is, every particular instance of heat is an instance of molecular motion, but there could be a situation in which there was some heat which was not molecular motion. (And similarly, with any of the other examples of contingent identity.) Is this plausible?

Two preliminary points. First, notice that this response does not contend (as my response does) that there could be something which felt like heat and was not molecular motion: but rather that there could be something which was the same (kind of thing) as heat, and yet not be molecular motion. I find Kripke's arguments in this context convincing. One can hardly imagine this contention to be true--unless we are implicitly interpreting it to mean just that something could <u>feel</u> like heat and not be molecular motion. For any situation in which there is no molecular motion is one in which there is no heat.[7]

Secondly, the very notion of a particular instance of heat used by (18') is quite murky. Is the instance of heat in the corner the same as that which fills the entire room? or a part of the room? or neither? Is it the same instance as that of a second ago? And when it comes to talk of different possible worlds, the problems are even worse.

---

[7]This point is more convincing when buttressed by the full force of Kripke's argumentation.

The crucial point is, however, this. What the CITs standardly asserted is that heat (the universal) is molecular motion, or that light is electromagnetic radiation, or that water is H_2O. If the CITs accepted (18') or something like it, in the case of other contingent identities, then they would again be vulnerable to Kripke's attack. For they would be admitting that heat is possibly not molecular motion; the second conjunct of (18') says that there is a possible world in which heat is not coextensive with molecular motion, and that surely is sufficient for the two universals to be possibly nonidentical. But then, by Kripke's argument, it follows that they are, in fact, not.

So (18') and its analogs (including (18)) is incompatible with CITs' claim that, e.g., heat is molecular energy. Consequently, the reductio of the CIT position by Kripke, can not be blocked by appeal to (18)--for if they did accept (18) and its analogs, they would be no better off.

2) So far, we have seen that (18) is not an acceptable premise for a CIT. There still remains the question of whether, independent of the views of the CITs, we could use (18) to refute Kripke; whether we could analyze our intuition that people might not have been bodies as the intuition that (18) holds, and thus, knock out the basis for (14) and Kripke's conclusion of nonidentity.

The discussion of this question will assume that bodies are essentially bodies. Those who would deny this

would, in any case, find appeal to (18) superfluous. (I shall discuss such a denial in point 3) below.) For them, not only might there be people without bodies, as (18) asserts, but even actual people, actual bodies--e.g., Descartes--might not have been bodies, and still (14) would not follow. So they find the intuition that possibly Descartes is not a body, quite conpatible with their materialism, and do not need (18) to explain away that intuition. Let us, therefore, in discussing (18), restrict ourselves to the assumption that bodies are essentially bodies. Given this assumption, can (18) be used to explain away our intuitions in such a way so as not to commit us to (14)?

Though there appears to be nothing formally wrong with (18), I think it unacceptable. For 1) on the most plausible assumptions, it is false; 2) it seems incompatible with our present assumption that bodies are essentially bodies; and 3) even if we disregard the above two points, (18) does not seem to capture the content of our intuitions.

First point: The CITs were, I believe, to a certain extent, correct. One should take the "person" case as analogous to the heat (lightning, water, etc.) case. Just as to be heat is to be a particular kind of thing, viz. molecular motion, and nothing which is not molecular motion is heat, so, too, being a person is being a kind of thing--and for the materialist, that should be being a body of a certain kind, and nothing which is not a body of this kind could be a person. We have it then that (18) is false, for it claims that there could be persons which were not bodies;

so much anyway on, what seems to me, the most plausible assumptions. Perhaps this is just a way of reiterating for 'person' the claim that Kripke and Putnam have made for terms like 'water', 'tiger', heat, etc.: that they behave like rigid designators.

Second point: If we grant that (18) is true, then it seems that one can not reasonably discount the possibility that actual people might not be bodies. Having a body is not a prerequisite for being a person, as (18) admits, why then might I not have been such an entity, though this is perhaps physically not possible? If persons might not be bodies, what metaphysical necessity demands that I, a person, must be one? The only possible reason would be that I am identical to my body and bodies are essentially bodies. But to advert to this is, to a certain extent, to beg the question.

Further, (18) and the claim that bodies are essentially bodies, seem jointly incompatible with a widely held view (one of the few) on the nature of essential properties. The view is, loosely put, that aside from such "trivial" properties as being self-identical (or being identical with Israel Krakowski), entities have the essential properties which every object in the natural kind to which they belong have and no others. On such a view, if some objects of a kind are essentially bodies, then every object of that kind must be so. But (18) claims that there can be

persons who are not bodies, and hence, not essentially bodies; though, on present assumptions, some persons--i.e., those which are bodies--are essentially bodies.[8]

To this last point, there is an obvious response Feldman might make. To be a person, he might respond, is not to be any particular kind of thing. 'Person' is not a "natural kind," or "sortal" term; it is rather a term such as, say, 'baseball player', or 'red object'. Thus, it is reasonable to suppose that all baseball players are persons (or for a materialist, bodies), but that there could be baseball

---

[8]The formulation of this point needs to be tightened up considerably, but this is not the place to do it. For a nice presentation of a view of this sort, see: Bernard Enc, "Necessary Properties and Linnean Essentialism," Canadian JP, Vol. V, No. 1.

One problem with my formulation is that it must be modified for the case in which we have a hierarchy of kinds. Hydrogen atoms, besides being essentially atoms, are also essentially Hydrogen atoms, while Carbon atoms are not only essentially atoms, but also essentially Carbon atoms. Similar points could be made with other hierarchies.

Could Feldman now respond to my point about essential properties by claiming that the person case is analogous to the above? Perhaps, embodied persons are essentially bodies as well as essentially persons, while possible disembodied people are essentially disembodied as well as being essentially people--analogously to the other hierarchies of kinds. First of all, Feldman would have trouble with this move since he has already rejected on behalf of the materialist the claim that (actual) people are essentially people (See Fn. 2).

In any case, there seems to be a crucial difference between the person case and the others, though I am hard put to articulate it precisely. It is the difference between the relation of lions, tigers, and even possible unicorns to the general class of animals on the one hand, and the relation between embodied people and disembodied people to people, on the other. The former stands as species to genus, and therefore, forms a hierarchy of kinds, while the latter does not. There would certainly be something very fishy about saying that in this latter case, too, we have a hierarchy of kinds.

players who were not persons (bodies) in some possible worlds. But this is surely not incompatible with the claim that all actual baseball players are essentially persons (bodies).

Being a person, Feldman would claim, is in this respect like being a baseball player; it is not to be any particular kind of thing. Hence, there is no incompatibility between (18) and the claim that bodies are essentially bodies.[9]

The trouble with this response, it seems to me, is just the claim that 'person' is not a sortal. I certainly do not have any criterion for distinguishing sortals from other terms, but whatever criterion one uses, it surely seems that 'person' is just as much a sortal term as any other.

Could there be a reason for denying that this is so? Is it somehow that persons are not basic to our conceptual scheme (whatever that precisely means) the way, say, bodies are? That hardly seems true: persons seem as fundamental as anything to our view of the world (one thinks naturally of Strawson in this context).

Is it because there are no scientific laws which range over persons? But what of economics, psychology, sociology? Whatever criterion one uses, 'person' seems to

---

[9]Perhaps this is just another way of saying that Feldman respond that 'person' does not act like a rigid designator.

qualify.[10] So it seems that my original point stands.

And in further support of this point, it seems easy enough to describe a situation in which I might not have been this or any body. Thus, imagine that I am transmuted, over a long period of time, into ghost stuff, retaining all along the memories, personality traits, etc., that I presently have.[11] I would still be a person; wouldn't I be the same person?[12] Perhaps, if one does not countenance

---

[10]The obvious exception to this is the view that the only sortal terms are the "basic" predicates of physics--e.g., 'collection of basic particles', or what not. This might be the view of someone who thought that the only things which "really" existed were those mentioned in our most basic physical theories.

If we did use this criterion, then 'body' would not be a sortal term either; nor would being a body be an essential property of bodies, contrary to our present assumptions.

Thus, from the point-of-view of physics, we would have to identify a body with a particular set of atoms say. (Or even better, an n-tuple of such sets, one for each instance of time in the person's existence.) But there are certainly possible worlds in which that precise set of atoms would not constitute a body. So the body--the set of atoms--is not essentially a body. This sort of argument needs some obvious tightening up, but the general point is, I think, clear.

[11]Perhaps, Feldman would respond that, in such cases, one still had the same body (ghost stuff is still stuff after all). Though we grant this, (18) still claims that there might be persons without bodies; so whatever that is supposed to amount to (I leave the lurid details to science fiction buffs), we are to imagine the present case as one in which I am transmuted into a bodiless person as construed by (18).

[12]Since, by hypothesis, I am being transmuted into whatever it is that (18) has in mind when it talks of bodiless people. Perhaps, such a transformation is physically impossible, but it should certainly be metaphysically possible if (18) is true.

the possibility of such bodiless entities altogether, the answer might reasonably be no. But if with (18), we do countenance such entities, the above scenario sounds quite plausible.[13] So again, if we accept (18), we should not, contra our present assumptions, accept both that persons are bodies and that bodies are essentially bodies.

Third point. Even if we disregard the above two points, we have still not captured, by (18), the intuition expressed when we say people might have have been bodies. For this intuition, whatever its merits, arises most naturally from one's own case; from the feeling, in my case, that I might have existed without a body. This is not just the intuition that there might have been bodiless entities, ghosts as it were, but the stronger intuition that actual people might not have had bodies.

We need not, of course, accept this intuition at face value; there surely may be ways of explaining it away; but it is that intuition which needs explaining away. This (18) has not done.

---

[13]There are two questions here. The first is what properties a person might lose or gain and continue to exist. This is what the transmutation story seems most relevant to; it is the question of "personal identity." When we speak of essential properties, however, we are concerned with 'the (timeless) properties an object could not have failed to have and the properties it could have lacked while still timelessly existing' (To quote Kripke).

What the above example should then be construed as showing is, given the metaphysical possibility, persons granted by (18), the metaphysical possibility that there be a time at which I--a person who has a body throughout his entire existence--do not have a body. This will suffice, I believe, to make my point.

The above three points are not demonstrative, though I do find them cumulatively convincing. Let us, however, recall the role we are presently attributing to (18). It is to allow that (14) might be false, granting that bodies are essentially bodies, while giving expression to the felt contingency of the relation between persons and their bodies. My point has been that (18) can not reasonably be said to have succeeded at this. That, therefore, the burden of proof remains with Feldman.

3. The discussion of the above point assumed that bodies are essentially bodies; without this assumption, appeals to (18) are superfluous.[14] I now turn to the question of the nonessentiality of bodies. The complementary question of whether the materialist can claim that persons=bodies are not essentially persons (in notation: $(x)(Px \supset (Bx \& \neg \Box Px))$, has already been raised by Feldman on the materialist's behalf (See Fn. 2). The claim that they are not essentially bodies, either, can also, claims Feldman, be put to good effect by the materialist. Let me note at the onset that this whole discussion assumes the coherency of essential attribution.

Feldman contends that even were a materialist to concede to Kripke his intuition that Descartes might not have been a body, (14) would still not follow and the

---

[14](18) is formally compatible with the claims that persons=bodies are not essentially persons, are not essentially bodies, neither, or both of these.

materialist still might contend that Descartes was identical with his body. For the contrary would follow only if being a body were an essential property of bodies; if it is not then a situation in which Descartes is not a body may merely be a situation in which Descartes' body is not a body. Compare: This man is identical with this baseball player, and necessarily so. Yet, this man might not have been a baseball player.

I do not think such a view can be demonstratively shown false. Still, if a materialist were to use it to defend his materialism, it would be, I feel, incumbent upon him to justify it. Given that we are granting the coherency of essential attribution, could it really be that a body might not be a body? that, say, Quine's body might not be a body and still be the same thing? Such a thesis is by my lights prima facie false: there is no reason to believe it and Feldman does not attempt to produce such a reason.

To be sure, <u>if</u> one insists that Descartes is identical with Descartes' body, <u>and</u> that Descartes might not have been embodied, it then follows that Descartes might not have been a body. But should we not rather make the inference the other way around, if at all? That nothing could possibly be Descartes' body which was not a body, and hence, if Descartes can exist without being embodied, then Descartes is not identical with his body?

What would one think of the following (analogous) argument for the contention that not all planks of wood are essentially planks of wood. This ship (assume I am pointing to a wooden ship) is identical with these planks of wood. This ship could have been, and still might be, made of metal (or ghost stuff or whatever); in which case it would not be planks of wood. Since the ship <u>is</u> these planks of wood we can conclude that these planks of wood might not have been planks of wood.

I take it <u>this</u> is not a reasonable argument. Aha! Any ship materialist would realize that his view entails that planks of wood are not essentially planks of wood (sayeth Feldman). But is this plausible? If this example shows anything, does it not rather show that these planks of wood are not identical with this ship? (This discussion would remain unaltered if instead of planks of wood, we had planks of wood in a certain organization.)

I think it safe to say that Feldman has not presented the materialist with a plausible position. All he has, again, done, is shown that Kripke has not proven his case; that if we deny certain assumptions the conclusion no longer follow. Of course. But he has not given good reasons for denying the assumptions.

Feldman's exposition may read more convincingly than it should because of the disputes that have revolved around the notion of essential attribution. But these have dealt

mainly with the coherence of the notion (which Feldman apparently concedes). Thus, it may seem that any attribution of an essential property is moot. But if the coherence of the notion is conceded, certain attributions are a lot more reasonable than others; and one such reasonable one is that bodies are essentially bodies.

4. This fourth, and last point, will be a brief digression to take up Feldman's other major point. As we have just seen, Feldman claims that Kripke must assume that bodies are essentially bodies; if not, then the possibility of Descartes not having a body does not entail the possibility of Descartes not being identical with his body. But if we grant this, claims Feldman, then we might just as well use Leibnitz's Law to get Kripke's conclusion, obviating the need for his argument. The Leibnitz's Law argument would be that Descartes≠Descartes' body, since Descartes has a property his body does not: possibly existing where no body does.

Granting the validity of this latter argument does not, however, make Kripke's argument superfluous. One may think what one will of essential properties, and yet have to deal with Kripke's argument; for in the formulation of that argument, he does not advert to essential properties (in defense of (14) or otherwise).

To be sure, Kripke does countenance essential properties and appeal to them in response to those who deny that bodies are essentially bodies or persons essentially persons. But doing so is not part of his original argument; it is only in response to an objection, and it is the objection which raises the issue of essential properties. Thus, we might eschew all talk of essential properties and still be faced with Kripke's argument.

Further, whether or not the Leibnitz's Law argument makes Kripke's argument superfluous or not, that the issue can even be put clearly in terms of Leibnitz's Law is, in part, due to Kripke's analysis. Prior to Kripke's argument, one might have heard a discussion such as this:

A: Descartes is not identical with Descartes' body, for Descartes has the property of possibly existing where no body does, while Descartes' body does not.

B (a CIT): Your argument can not be valid. If it were, we could also show that heat≠ molecular motion; for heat has the property of possibly existing where no molecular motion exists and molecular motion does not. But heat <u>is</u> molecular motion. So the argument is invalid.

The virtue of Kripke's argument is that it helps disentangle these issues and see why B's response will not do.

<u>The Pain State/Brain State Argument</u>

Kripke's argument against brain state/pain state

identity is formally analogous to his argument against person/body identity: Let 'p' be a rigid designator of some particular pain and let 'b' be a rigid designator of the putatively identical brain state. If b=p, then □b=p; but it is possible that b≠p; therefore b≠p.

Feldman again attacks the premise that possibly b≠p. For though we concede that there might be a world in which this very brain state was not a pain state, it does not follow, claims Feldman, that in such a world, this brain state is not identical with this pain state; for such a world might merely be one in which this pain state is not a pain. Only if pains are essentially painful would it follow that in such a world, b≠p. But Kripke has not shown this to be so, and his conclusion does not, therefore, follow.

Kripke is aware of this response and says (as quoted by Feldman):

> The difficulty can hardly be evaded by arguing that being a pain is merely a contingent property of a, and that, therefore, the presence of b without pain does not imply the presence of b without a. Can any case of essence be more obvious than the fact that being a pain is a necessary property of each pain? Consider a particular pain or other sensation that you once had. Do you find it at all plausible that that very sensation could have existed without being a sensation, the way a certain inventor (Franklin) could have existed without being an inventor?[15]

---

[15]Kripke, Pg. 335.

Feldman's response is that:

> Any serious materialist should recognize that his view entails that painfulness is never part of the essence of a pain-event. Pain events are experienced as they are only as a result of the contingent laws of nature . . . these very pain events, had the laws of nature been different, would of course still have been self-identical, but would not have been identical to anything that would, under those circumstances, have been a pain event. Thus, such events are not essentially painful. The very same point can be put more straightforwardly by saying that certain brain events are such that it is contingent that they are felt as pains.[16]

If, then, it is open to one to reject the assumption, as Feldman does, that there are essentially phenomenological events--e.g., that pains are essentially painful--it seems the best Kripke can hope for is a stalemate.

Another argument mentioned by Kripke, though he does not go into detail, is that this very pain could have existed without being a brain state. Feldman suggests that the materialist respond to this by denying that brain states are essentially brain states.[17]

---

[16] Feldman, Op. Cit., Pg. 675.

[17] Another way the materialist might deal with this is by appeal to an analog of (18). All arguments made about (18) would then go over to this case. E.g., such a view would not do for a CIT since it would entail that pain (the state) is not identical to CFS, since it would admit the possible noncoextensivity of the two.
Since Feldman does not mention this possibility, I mention it only here. Similarly, for the contention in the text that brain states are only contingently brain states.

Feldman's contention that the materialist can deny that pains are essentially pains has, in particular, received wide attention; I shall, therefore, examine it in some detail below. Before doing so, however, let me briefly note that a Feldman-type response does not work against Kripke's type/type argument. (Feldman does not say that it does.)

Particular pains, we are assuming, have the contingent property of being felt as painful. But what of that property? Given any physical property (of brain events, their structure for instance) it seems possible for that physical property to obtain and there be nothing felt as pain; that is, without the contingent property of being felt as painful obtaining. Since the properties are then possibly not identical, they are, by Kripke's argument, actually not. So this (mental) property is not identical with any physical property.[18]

To return to token/token identity. It appears that Kripke accepts Feldman's inference for, as we saw, he takes it that the claim that pains are essentially painful, is crucial to his argument (and also obviously true). If so, he does himself a disservice: he can, in fact, concede (for argument's sake) that pains are not essentially painful without doing his cause damage.

---

[18]Since Feldman does not consider type/type identity, I shall not pursue this line of thought in detail.

Intuitively put, my point is this. Kripke focuses on the phenomenological quality of mental events. Conceding Feldman's point will not, appearances to the contrary, prevent his doing so; viz., instead of considering a particular pain, Kripke could now consider the particular event of that pain's having the contingent property of being painful. This event, if physicalism is true, must also be physical; but Kripke's argument will now apply to it.

Feldman's claim--just quoted--is that it is only because of the contingent laws of nature that pains are experienced as they are. Consider now some one particular pain (you once had), 'Abe' let us call it. This event, Feldman asserts, has the contingent property of being experienced as pain(ful) (we may call this contingent property 'Q'),[19] just as a baseball player has some contingent property which makes him a baseball player--the contingency of which entails that this baseball player might not have been a baseball player.

But Feldman's move does not really avoid Kripke's conclusion: it merely postpones it. Instead of discussing pains, and Abe, in particular, we can now discuss the particular event of Abe's being experienced as pain, or put another

---

[19] Feldman, as quoted above, talks as if a pain's being a pain is identical to a pain's being experienced as pain(ful). From which I suppose we can conclude that he believes that the property of being a pain is identical with the property of being experienced as pain(ful). In any case, I shall let the contingent property Q be the property of being experienced as painful in this note.

way, Abe's having the contingent property Q. Let us call this event Issac.

Kripke could now argue that Issac is not identical with any physical event; for given any physical event to which Issac is purportedly identical (an 'IFS' we might call it, on analogy with 'CFS') we could imagine that physical event occurring without anything being felt as pain, that is, without Issac occurring.[20] By Kripke's argument, Issac is then not identical with the IFS, while a physicalist is surely committed to Issac's being physical (since he is committed to all events being physical).

Here I am assuming, of course, that there is such an event as Issac, an assumption based on the (popular view) view that for any individual i, and for any property of i, P, there is an event or state of i's having P.[21] And all my argument requires is the noncontroversial part of the above analysis--i.e., that the above gives a sufficient condition for eventhood. It would be a very strange physicalism indeed which <u>required</u> as one of its premises the denial of this sufficient condition. And such a denial would be of little use to Feldman. For he argues that "certain brain events are such that it is contingent that they are

---

[20]This intuition is certainly something the Contingent Identity Theorists would agree to. It is, on their view, only due to the contingent laws of nature that Neurophysiological events are experienced as they are.

[21]The most prolific proponent of this view is Jaegwon Kim; see, e.g., "Events as Property Exemplifications," in <u>Action Theory</u>, Myles Brand and Douglas Walton, eds. (Dordrecht, Holland, Reidel, 1976), pp. 159-177.

felt as pain." Given this and Feldman's countenancing of the existence of events, he is hardly in a position to deny the existence of Issac.

Kripke's argument depends on the claim that, " . . . if the identity thesis were correct, the element of contingency can not lie, as in the case of heat and molecular motion, in the relation between the phenomena (heat = molecular motion) and the way it is felt or appears (sensation S), since in the case of mental phenomena, there is no appearance beyond the mental event, itself." This is what Feldman denies: The mental phenomenon does have an appearance beyond it; besides the phenomenon itself, there is the way it contingently happens to feel. But this, I have argued, merely pushes off the problem from the mental phenomenon to the mental phenomenon's feeling a certain way, from Abe to Issac.

My point so far has been that Feldman's response sticks him with an event--Issac--which a Kripke-style argument shows not to be physical. But perhaps Feldman could now defend the physicality of Issac analogously to his defense of Abe's being identical to some particular CFS.

The defense might be that the imagined possibility of the particularly IFS occurring without anything being felt as pain is due to the fact that it is only a contingent property of the IFS (=Issac) that it is an experiencing of

something as pain. What we are imagining is not the IFS occurring without Issac, but simply Issac (= the IFS) not being an experiencing of something as pain.

But again, this sort of move will not work. First, recall that Issac is the event of Abe's being experienced as pain; now, could the event of Abe's being experienced as pain not be an experiencing of something as pain? The chair I am sitting on is black and contingently so; perhaps, it is even contingently a chair. But could the state or event of this chair's being black (or being a chair) possibly exist without being a state or event of something's being black (or being a chair)? Such a view is doubtful to say the least, and so, too, the analogous claim about Issac.[22] It is even more dubious than the idea that pains might not be painful.

Unintuitive or not, Feldman could, I suppose, stick to his guns: Issac, he might say, a particular physical event, is only a case of something's being felt as pain because of the contingent laws of nature. There is some contingent property (C) which it has in virtue of which it is an event of being experienced as pain. If this were right, a Kripke-type argument would still not work.

---

[22]The problem here is not the ascription of contingent properties to events. One can say that the assassination at Sarajevo contingently has the property of causing WWI. What would be implausible would be to say that the assassination at Sarajevo's causing WWI might have not consisted in something's causing WWI.

But, if Feldman would thus use this same sort of response in the Issac case, we could make, pari parsu, the equivalent counterresponse; there is now a certain further event, i.e., Issac's having the contingent property C--let us call this event 'Jacob'--which we can, in an analogous manner, now argue is not physical. Either we stop positing, at some level, these strange contingent properties, at which point we can apply Kripke's argument, or we are in an infinite regress: Abe, Issac, Jacob . . . .[23]

And this would, I believe, be a vicious infinite regress. It would entail the existence, in one person, of an infinite number of contingently related neurophysiological events. But a person's functioning brain and nervous system exists for but a finite amount of time and have only a finite number of, say, basic particles. If we assume that there is a non-infinitesimal lower bound on the duration of a neurophysiological event and that each such event must involve at least one basic particle, then it seems clear that there could not be the required infinity of contingently related events.

---

[23]Or more awkwardly: Some particular pain, that pain's being experienced as pain, (that pain's being experienced as pain)'s being an event of something's being experienced as pain. Probably, the most perspicuous way of putting the sequence (due to Ned Block) is:

1. The pain.
2. 1's being experienced as pain.
3. 2's being an event of something's being experienced as pain.

Could, however, Feldman circumvent the viciousness of the regress by claiming that the above argument is valid only if we assume the events in question (Abraham, Issac, Jacob . . .) are distinct contingently related events. That if we assume that these events are logically, semantically, or set theoretically related, then the existence of the infinite series of events: Abraham, Issac, Jacob . . . , is no more strange than the existence of the infinite series: my finger, the set containing my finger, the set containing the set containing my finger . . . .[24]

For Feldman, such a reply can not work; it is incompatible, I believe, with his requirement that the particular events (Abe, Issac, Jacob) have their corresponding properties only contingently.

Consider, as an example, the claim that each of the events is related to the others by the relation of identity. That is, Abe, Issac, Jacob . . . are all identical to one another and to some one neutral event--i.e., some particular CFS. This would be an attractive line for one who thinks, as Feldman does, that pains <u>are</u> experiencing as pains.

But recall that it is supposed by Feldman that Abe can occur without being felt as pain; so Abe can occur

---

[24] One recalls Bradley's generation of an infinite number of relations given just one of the form aRb; to which I have heard it very sensibly replied, "So what?"

without Abe's being felt as pain.[25] Then Abe can exist without Issac existing: so Abe is not identical with Issac, contra hypothesis.

So this sort of reply could not be used to prevent Feldman's impalement on the horn of infinite regress. And similar problems will attend any such attempt to avoid the viciousness of the regress by taking the relation between the events, Abraham, Issac, Jacob . . ., to be some (other) logical, semantical, or set theoretic relation.

In sum, Feldman's position leads to a dilemma: since it must allow for the existence of Issac distinct from Abe, we can ask whether Issac is essentially an experiencing of something as pain (as is most plausible) or not. If it is, then we can proceed to apply Kripke's argument to Issac. If, on the other hand, it is not, we are caught in a vicious infinite regress.

And this dilemna is not an artifact of my interpretation of Feldman. It is intrinsic to his argument: the offspring of his claim that for mental phenomena--e.g., pain--there are appearances beyond the phenomena itself--i.e., they only contingently feel the way they do. For we may now enquire after the status of these further appearances.

---

[25]The validity of the argument depends on the validity of the claim that where Θ is a term for a property, $\ulcorner$a's having (or being) $\Theta\urcorner$ obtains if a has (or is) Θ. This claim would be very hard to deny.

CONCLUSION

Conclusion

This thesis has examined the views of two philosophers. The result of this examination is that (especially) type/type materialism must, given the present state of philosophical discussion, as represented in this thesis, be tentatively considered false. The basis for this judgment is that (1) the most cogent defense of this view I know of, Lewis', has major flaws; and (2) there remains unanswered a forceful challenge to materialism in general, Kripke's.

The flaws discerned in Lewis' views are varied. First, we noticed that his relativization to context of everyday mental terms engenders problems. There are two ways of analyzing this move. The first, while preserving the type/type character of Lewis' claims, applies only to a limited domain, and is empirically falsifiable. The second, while of unrestricted domain and immune from falsification, can no longer properly be considered a type/type view. Both analyses run into trouble when we ask whether they should be extended to the more general case of theoretical terms.

We have also seen that Lewis defines 'pain' so that it refers to a physical entity, but does not so define 'the property of having pain'. Doing this engenders the

problem, along with some other technical ones, that the most crucial (it seems to me) property for a type/type materialist to reduce, has not been so reduced. Similar problems are seen to obtain for Lewis' more general theoretical considerations. Lewis does not define the kinds over which the laws of special sciences make true, interesting generalizations, so that they come out physical.

Further, when discussing Kripke, we have seen that his antimaterialist claims remain unanswered. Dummett's line of attack is that both Kripke's metaphysical/epistemic and rigid/nonrigid distinctions, are properly construed as distinctions of scope; and that, there is, consequently, no distinction to be discerned between proper names and descriptions. Kripke's arguments do not fall prey to this line of attack.

Chomsky attacks a different aspect of Kripke's views. His main argument is that Kripke's notion of essential attribution reduces to that of de dicto necessity. We concluded that Chomsky's argument has no force against one who is already committed to essentialism.

Finally, Feldman's critique aims directly at Kripke's arguments against the claims that, each person is identical with his or her body, and that each particular mental event or state is identical with some corresponding physical state or event. In the person/body case, his main thrust

is that we can account for the contingency felt to obtain between persons and their bodies in a manner innocuous to materialism. I showed that Feldman's analysis of this felt contingency is not compatible with the other views of contingent identity theorists, and that generally it will not do in its support of materialism.

In the mental/brain state case, Feldman argues that Kripke assumes and depends on the claim that mental states are essentially mental, e.g., that pains are essentially painful. An assumption that the materialist need not and should not make. It was seen that this view of Feldman's leads to a dilemma, either horn of which impales.

So, all in all, Kripke's arguments have remained unscathed.