
THE INTERNAL BASIS OF MEANING

BY

MICHAEL MCKINSEY

Recent work in the philosophy of language by Donnellan (1970, 1974), Kaplan (1968, 1979), Kripke (1972) and Putnam (1970, 1975) has, I believe, convinced many philosophers that there is something fundamentally wrong with the received, traditional point of view concerning the relation between language and mind. According to this received view, the semantic facts about the meaning of words in public languages are ultimately determined by facts about the internal psychological states of individual speakers of the language.¹ I will call the received view 'internalism' and its denial 'externalism.'

The most direct and powerful attack on internalism is found in Putnam's (1975) justly famous "Twin Earth" argument. From this argument, Putnam concludes that the meanings of at least one kind of word—natural kind terms—are not determined by speakers' "narrow" psychological states, where a psychological state is said by Putnam to be narrow only if it presupposes the existence of no contingent object besides the person who is in the state (1975, p. 136).

I believe that in one natural and important sense of 'narrow psychological state,' a sense that I will explain below, Putnam's argument pretty clearly proves its conclusion. However, Putnam apparently goes on to infer that, since the meanings of natural kind terms are not determined by psychological states that are narrow in his sense, these meanings must instead be determined by external facts concerning the historical, causal, and social relations that speakers bear to each other and to other objects in their environments. And it is this further inference, whose conclusion roughly characterizes a "causal theory" of meaning, that I wish to disagree with.

For there is a second natural sense of 'narrow psychological state,' and Putnam's argument fails to show that the meanings of natural kind terms are not determined by states that are narrow in this sense. The sense in

question is one on which *first person* or *de se* cognitive attitudes are allowed to qualify as narrow. I will try to make it plausible that speakers' first person cognitive attitudes provide the best explanation of how the meanings of natural kind terms are understood and determined in natural language. If this is right, then Putnam's argument cannot be used to show that internalism is false. For, or so I will argue, the facts about speakers' first person attitudes that determine the meanings of natural kind terms are *internal* facts, facts that are logically independent of any facts about the external (causal, historical, social) relations that the speakers may bear to other speakers or objects.

In Part I, I will attempt to clarify two senses in which a psychological state can be said to be narrow. These two senses will in turn provide two different versions of internalism. The upshot of this part will be that *if* first person cognitive attitudes can be used to explain how the meanings of natural kind terms are understood and determined, then one of these versions of internalism escapes Putnam's argument unscathed.

In Part II, I will try to provide an explanation of how the meanings of natural kind terms can be grasped and determined by first person states. Part of the *prima facie* difficulty in giving such an explanation arises from the fact that the meanings of such terms are public and external (in a sense to be explained), while first person attitudes are private and internal. How can such meanings be grasped in such states? My answer to this question will be based on an argument I have given elsewhere (McKinsey, 1986) for the conclusion that, contrary to common opinion, cognitive attitude states are not essentially characterized by their propositional contents.

Part I. Two Forms of Narrowness

1. PUTNAM'S USE OF TWIN EARTH

In Putnam's well-known example, we are to imagine that somewhere in the galaxy there is a planet that is nearly an exact duplicate of Earth. This planet, "Twin Earth," has the same history and appearance as Earth, and the people and ordinary objects that exist on Twin Earth are identical copies of corresponding people and objects that exist on Earth. Thus each English speaker on Earth has a doppelgänger on Twin Earth who also speaks English and who duplicates his neurophysiology, personal history, and mental life.

The only major difference between Earth and Twin Earth is that the liquid called 'water' by English speakers on Twin Earth does not have the same chemistry as water. Instead of being composed of H₂O molecules, the corresponding liquid on Twin Earth is composed of, say, XYZ molecules. But aside from this chemical difference, the two liquids are

indistinguishable: both fill lakes, rivers, and oceans; both fall from the sky as rain; both are thirst-quenching, colorless, odorless, etc.

Now it seems clear that, as Putnam points out, the word 'water,' as we use it on earth, would have a different extension than the word as used by speakers on Twin Earth. For water is H_2O ; that is, all stuff that is water is composed of H_2O molecules. And so, only such stuff is in the extension of 'water' as we use it. This seems true not only as a matter of fact, but also in counterfactual situations: only stuff composed of H_2O molecules *would* or *could* be water. Thus, what the Twin Earthians call 'water' would not be water, since what they call 'water' is composed of XYZ, not H_2O .

But then, since a general term's extension is determined by its meaning, it follows that the word 'water,' as used by us on Earth, must have a different meaning than it would have when used by our counterparts on Twin Earth.² One serious problem raised by Putnam's example is that of explicating the difference between these two meanings of 'water'. Given the extreme qualitative similarities of water and XYZ, we must suppose that ordinary Earthians and their Twin Earthian counterparts would associate the same qualitative characteristics with the word 'water,' and so we cannot appeal to such characteristics to account for the difference in meaning. Nor can the chemical differences between water and XYZ provide a solution to our semantic problem. For the different chemical characteristics of water and XYZ would be unknown to most ordinary Earthian and Twin Earthian speakers. Thus these chemical characteristics also cannot be used to account for the difference in the meanings of 'water' as used and understood by ordinary (scientifically ignorant) Earthian and Twin Earthian speakers.

I have proposed a solution to this semantic problem elsewhere (McKinsey, 1987), and I will return to a discussion of it below. But Putnam himself emphasizes a different, though closely related, difficulty. This is the problem of accounting for the psychological difference between Earthians' and Twin Earthians' *understanding* or *grasp* of the two meanings of 'water.' That there must be such a difference in understanding follows from the fact that Earthian and Twin Earthian speakers who correctly understand the meanings of 'water' in their respective dialects of English must be understanding or grasping two different meanings.

It is perhaps the main contention of Putnam's discussion that the difference in Earthians' and Twin Earthians' understandings of the meaning of 'water' cannot be accounted for by appeal to differences in these speakers' *narrow* psychological states. Consider an English speaking inhabitant of Earth, Oscar, and his Twin Earthian doppelgänger Toscar. Suppose that Oscar and Toscar, being competent speakers, both perfectly understand the meanings of 'water' in their respective dialects.

Also assume that neither knows anything about the chemistry of the different liquids called 'water' on their respective planets (perhaps, Putnam suggests, Oscar and Toscar live prior to the development of chemistry). Thus Oscar and Toscar would associate exactly the same qualitative characteristics with the word 'water.' Moreover, we may suppose, Oscar and Toscar are "exact duplicates in appearance, feelings, thoughts, interior monologue, etc." (Putnam, 1975, p. 141.) Oscar and Toscar would therefore share, according to Putnam, all the same narrow psychological states. And yet their understandings of the meaning of 'water' would be different, since each understands 'water' to have a different meaning than the other understands. Thus, Putnam concludes, what a person understands a word's meaning to be, what in effect a person *means* by a given word, is not determined by that person's narrow psychological states.

It seems reasonable to take Putnam's argument as directed against the following principle, which I will call "meaning-internalism":

- (MI) Necessarily, if two persons share all the same narrow psychological states, then one of the persons means something by a given word if and only if the other person means the same thing by that word.

(MI) is one of the internalist principles that I wish to defend below. More exactly, I will defend a version of (MI) in which 'narrow psychological state' is understood in a certain sense, a sense on which first person cognitive attitudes can qualify as narrow. (MI) provides one plausible interpretation of the claim that a person's linguistic understanding is *determined* by that person's narrow psychological states. This seems in fact to be the same sense of 'determines' that Putnam himself uses, as when he equates the principle that intension *determines* extension with the principle that "sameness of intension entails sameness of extension" (1975, p. 136).

2. FIRST PERSON STATES AND PUTNAM'S ARGUMENT

It is important to note that when we speak of two persons sharing the same psychological states, we do not mean to be talking about particular or concrete states of the person. For such states, like particular concrete events, are not sharable. Thus, Oscar's believing at a time *t* that birds fly is a concrete state of Oscar that is numerically distinct from Toscar's believing at *t* that birds fly. The sense in which Oscar and Toscar share the same psychological state in this case is that they are both in the same *type* of state: both Oscar's and Toscar's particular states of belief are states of believing that birds fly. Putnam speaks of psychological states

as "predicates" (1975, p. 136). In the same vein, I will speak of them as *properties*. Thus on this understanding, two persons share all the same narrow psychological states just in case they share all the same narrow psychological properties.

One of the crucial premises of Putnam's argument is that a typical English speaker and his doppelgänger on Twin Earth would share all the same narrow psychological states. In particular, Putnam assumes that the pairs of twins in his example would be "exact duplicates in appearance, feelings, thoughts, interior monologue, etc." (1975, p. 141). But is this assumption really plausible?

Certainly the assumption seems false with respect to one large and important group of psychological states that might naturally be classified as "narrow." I have in mind *first person* cognitive attitudes, such as beliefs of the kind ascribed by sentences of the form 'x believes that he himself is F.'³ Suppose that the following is true of Ollie, a militaristic and self-aggrandizing inhabitant of Earth:

(1) Ollie believes that he himself is a hero.

Would Tollie, Ollie's counterpart on Twin Earth, share the belief that (1) ascribes to Ollie? Apparently not. For Ollie's belief is about *Ollie*, and we may assume that Tollie has no beliefs about Ollie at all. To be sure, Tollie has a very *similar* belief, for Tollie no doubt believes that *he* himself is a hero. But this latter belief, being about Tollie, is not the same as Ollie's belief. Thus Ollie and Tollie have different first person beliefs, so that if such beliefs qualify as narrow psychological states, then contrary to Putnam's assumption, Ollie and Tollie would *not* share all the same narrow psychological states.

It is easy to be confused about this. For the two sentences (1) and

(2) Tollie believes that he himself is a hero

may easily seem to ascribe the *same* psychological property to Ollie and Tollie. And so they do. The property in question is simply the narrow psychological property that any person x has if and only if x believes that he himself is a hero. (Cf. Lewis (1979), p. 142.) And in general, every true first person belief sentence concerning Ollie ascribes to Ollie a narrow psychological property that would be shared by Ollie's doppelgänger on Twin Earth. From this fact, one might infer that first person belief ascriptions can pose no problem for Putnam's claim that a competent speaker such as Ollie and his doppelgänger would share all the same narrow psychological states.

But this inference would be mistaken. Let H be the property that any person x has if and only if x believes that he himself is a hero. Then as

we've seen, (1) ascribes H to Ollie and (2) ascribes H to Tollie. However, there is *another* property that (1) ascribes to Ollie that (2) does *not* ascribe to Tollie. This is the property of having a certain self-referential belief, a belief about Ollie to the effect that he is a hero. Given that this property is narrow, then since Ollie has the belief in question and Tollie does not, Ollie and Tollie would not share all the same narrow psychological states, even though it is true that they both have the narrow property H.⁴ We may conclude, then, that Putnam's argument does not succeed against any version of meaning-internalism in which 'narrow psychological state' is understood in such a way that first person beliefs are allowed to qualify as narrow psychological states.⁵

3. WHAT IS A "NARROW PSYCHOLOGICAL STATE"?

At this point, a defender of Putnam's argument might object by insisting that first person beliefs should *not* be allowed to qualify as narrow psychological states. However, it is difficult to evaluate this objection, since the concept of a narrow state has not yet been made clear enough to allow us to tell definitely whether first person states should be classified as narrow or not.

In all of the literature in which the concept of a narrow state has been invoked, I know of no explanation of the concept that is any clearer than Putnam's initial statement, which says merely that no narrow psychological state "presupposes the existence of any individual other than the subject to whom that state is ascribed" (1975, p. 136). Now this statement might certainly lead one to classify first person beliefs as narrow states. After all, when we ascribe a first person belief to someone, our ascription would typically seem to presuppose the existence of no individual besides the believer in question. Thus, when we say that Ollie believes that he himself is a hero, it seems that only *Ollie's* existence is presupposed.

Moreover, it could be argued, first person beliefs should be considered virtual paradigms of narrow psychological states. After all, it is just such beliefs as Descartes' beliefs that he himself is a thinking thing and that he himself exists, which form the bulwark of the traditional Cartesian "methodological solipsism" to which Putnam is opposed.

On the other hand, it is certainly *possible* to construe Putnam's criterion in such a way that it rules out first person states as narrow. For instance, it is not implausible to suppose that Putnam had in mind a definition such as

- (3) The psychological property of being S is *narrow* if and only if: necessarily, for any person x, x's being S does not entail the existence of any contingent object other than x,

where in turn the *definiens* of (3) may be rendered as

$$(4) \quad \Box(\forall x) (\forall y) [(Cy \ \& \ \Box(Sx \supset Ey)) \supset x = y].$$

(Here 'Cy' and 'Ey' mean 'y is contingent' and 'y exists,' respectively. The box '□' means 'it is necessary that' and entailment is understood as necessity of the material conditional.)

It turns out that, on this construal of narrowness, first person beliefs are not narrow. This is due to a peculiarity of such beliefs that I have already emphasized. Thus consider again the belief ascribed by (1) 'Ollie believes that he himself is a hero.' No one could possibly have the same belief as that ascribed by (1) without having a belief about Ollie. But this means that someone other than Ollie, Ron say, could also not have the belief ascribed by (1) unless Ollie exists. Ron's having the belief in question, therefore, entails that Ollie exists. By criterion (3), then, first person beliefs cannot be narrow psychological states.

Now it is true of course that Ron *could not possibly have* the belief ascribed by (1). For not only must this belief be about Ollie, but it must also be a *self-referential* belief about Ollie. And of course no one but Ollie could have a self-referential belief about Ollie.⁶ But the point that Ron could not possibly have the belief ascribed by (1), while correct, only serves to reinforce my claim that Ron's having this belief entails that Ollie exists (since all propositions are entailed by impossible ones).

Thus a defender of Putnam's argument might appeal to the definition (3) to justify the claim that first person beliefs should not be classified as narrow. However, there is a serious difficulty with the idea that (3) captures Putnam's notion. For if narrowness is construed as (3) suggests, there would probably be no such thing as a narrow psychological state.

Consider any state that would normally be proposed as a paradigm of narrowness, such as my belief that there are philosophers. The proposition that I believe that there are philosophers entails whatever the proposition that I exist entails, since the former proposition entails the latter. But it may well be (if materialism is true) that my existence entails the existence of my body, or perhaps entails the existence of certain sperm and egg cells. (See Kripke, 1972, p. 314.) If so, then my believing that there are philosophers entails the existence of my body and of certain tiny cells. By (3)—with its *definiens* understood as (4)—it follows that my belief that there are philosophers (indeed, *any* belief of mine) would fail to be a narrow psychological state.⁷ This is not a happy result.

So there is a serious problem as to what Putnam had in mind when he introduced the notion of a "narrow" psychological state. Let us reconsider the intuitions that may have motivated this introduction. Intuitively, one might say, a narrow psychological state or property is one that is purely conceptual or purely qualitative. Thus a narrow state

should *not* be a state that is *relational* with respect to some ordinary, concrete, material object. By saying that a state or property is relational, I mean that it essentially involves some particular object. Examples would be such properties as those of admiring Ollie, living in Detroit, being shorter than Kareem, etc. Since we are only concerned with properties that are relational with respect to contingent objects, let us define

- (5) The property of being S is *contingently relational* if and only if: there exists some contingent object y such that necessarily, any object has S only if y exists (that is, $(\exists y) (Cy \ \& \ \Box(\forall x) (Sx \supset Ey))$)).

Perhaps Putnam wished to characterize a narrow psychological state simply as a psychological state that is not contingently relational. Such a definition might seem congenial to a defender of Putnam's argument, since it has the consequence that no first person belief is narrow. This is because, as we've seen, all first person beliefs are contingently relational (with respect to the believer). However, the definition has the serious drawback of ruling *in* as narrow all psychological states that *fail* to be contingently relational. But some psychological states that are not contingently relational seem nevertheless to be paradigm cases of states that should also not be classified as narrow. I have in mind properties that are existential generalizations of wide, contingently relational psychological properties, generalizations such as the property of thinking while under a tree or the property of having a belief about something that is in fact a rock. Since there is no *particular* object whose existence is implied by having either of these properties, both properties fail to be contingently relational.⁸ Thus the definition under consideration incorrectly classifies such properties as narrow, even though having them entails the existence of such external things as trees and rocks.

A correct definition of 'narrow psychological state' should not just rule out contingently relational states; it should also rule out states that, while they are not contingently relational, are states that nevertheless guarantee that *there exist* contingent objects external to agents that are in the states. This second requirement is difficult to spell out clearly. At first glance, it seems to mean that our definition should imply that a property S is narrow only if it is not necessarily guaranteed by any person's having S that there exist any contingent objects other than that person. Here, the condition we are considering seems best rendered as: $\sim \Box(\forall x) (Sx \supset \Box(\exists y) (Cy \ \& \ \sim(y=x)))$, or equivalently, as

- (6) $\Diamond(\exists x) (Sx \ \& \ (\forall y) (Cy \supset y=x))$.

We now have still another way of interpreting Putnam's requirement

that no narrow state “presupposes the existence of any individual other than the subject to whom that state is ascribed.” However, this interpretation also has some severe drawbacks. For one thing, as (6) makes clear, a psychological property *S* would be narrow on this interpretation only if it is possible for a person to have *S* even though *no other contingent object exists at all*. But this requirement is so strong that it seems highly doubtful that any psychological property could actually satisfy it. Is there really any psychological property such that an individual could have that property while, in effect, existing alone in the universe? Perhaps, for all we know, a certain level of complexity must be achieved by any being, in order for that being to be capable of having psychological properties; and perhaps that level of complexity inevitably requires the existence of other, perhaps simpler, contingent objects. If so, then no being is capable of both having psychological properties and existing alone in the universe. Surely, a definition of ‘narrow psychological property’ that presupposes otherwise is at best undesirably contentious.

Perhaps it might be suggested that a narrow psychological property should not be capable of being possessed independently of *all* other objects; rather, it should only be capable of being possessed independently of all *physical* objects. This suggestion would have us replace (6) by

$$(7) \quad \diamond (\exists x) (Sx \ \& \ \sim (\exists y)Py),$$

where ‘*Py*’ abbreviates ‘*y* is a physical object.’ But the suggestion that there are narrow psychological properties satisfying this condition also seems undesirably contentious.⁹ Consider the property of having a visual experience. Does the hypothesis that this property is narrow really commit us to its being possible for a non-physical (presumably “purely spiritual”) being in a non-physical world to have visual experiences? It seems to me highly doubtful that it is possible for there to exist “purely spiritual” beings that are capable of having psychological properties. But it is surely even more doubtful that such beings could have visual experiences, even though—being “purely spiritual”—they have no sense organs or nervous system, and even though the world they inhabit contains no (physical) light.¹⁰

4. CONCEPTUAL IMPLICATION

We seem to be on the wrong track. We have been trying to capture the idea that a psychological property is narrow only if an agent’s having that property does not “presuppose” that there exist contingent objects external to the agent. And we have been trying to capture the sense in

which narrow properties do not "presuppose" the existence of external objects by using the metaphysical notions of necessity and possibility. The repeated failure of our various attempts to apply this strategy strongly suggests, it seems to me, that these metaphysical notions will not serve our purpose.

One plausible hypothesis is that the concept we have in mind when we say that a narrow psychological state does not "presuppose" the existence of any physical objects, is an *epistemological*, rather than a metaphysical, concept. If this is correct, then it is important, for it means that the concept of a narrow psychological state is also at bottom an epistemological, rather than a metaphysical, concept.

Consider the original context that motivated the introduction of the idea of a narrow psychological state, namely, the Cartesian predicament. Descartes' problem was that, although he knew with certainty that he existed and that he had ever so many psychological characteristics, he was unable to correctly *deduce* from this certain knowledge that an external world exists or even that he himself had a physical body. But notice that this failure of deduction does *not* show that it is metaphysically possible for Descartes to exist as a disembodied mind in a non-physical universe. Rather it shows that, *for all one can deduce* from the facts about Descartes' narrow states, Descartes has no body and no physical objects exist. Or perhaps we could say that, relative to Descartes' knowledge of his own narrow psychological states, it is *epistemically possible* that no physical objects exist.¹¹

Let us say that a proposition *p* *conceptually implies* a proposition *q* if and only if there is a correct deduction of *q* from *p*, a deduction whose only premises other than *p* are necessary or conceptual truths that are knowable *a priori*.¹² What Descartes discovered was that the facts he knew incorrigibly about himself and his own mental states do not conceptually imply the existence of a physical world. But notice that it is consistent with this discovery to suppose that, say, Descartes' existence entails that he has a physical body, and that Descartes necessarily originated from such-and-such sperm and egg cells. For these latter propositions, if they are true at all, seem to be knowable only *a posteriori*, and so these propositions are not conceptually implied by Descartes' knowledge of his own mental states.

Perhaps, then, narrow psychological states do not "presuppose" that there are physical objects in the sense that such states do not conceptually imply that there are physical objects. And also, perhaps, we should define a psychological property *S* as being *narrow* if and only if: (1) *S* is not contingently relational, and (2) for no object *x* does the assumption that *x* has *S* conceptually imply that there exist physical objects.

This is close, but not quite correct. For intuitively, the idea that a state is narrow does not mean *merely* that one cannot correctly deduce the

existence of physical objects from someone's being in the state. Rather, it means more generally that one cannot correctly deduce the existence of anything outside the agent's immediate mental experience.¹³ Consider, for instance, the property of having had a thought precisely two thousand years after someone else had the same thought. Clearly, a person's having this property does not conceptually imply that there are any physical objects. Yet the property should not count as narrow, since one can deduce *a priori* from one's having the property that someone else existed two thousand years in the past.

Given this general consideration, then, perhaps we should define a psychological property S as being *narrow* if and only if: (1) S is not contingently relational, and (2) for no object x does the assumption that x is S conceptually imply either that (a) there exist physical objects, or that (b) there exist contingent objects other than x and x's own mental states, acts, and experiences.

Now while I think that this condition is certainly *necessary* for a psychological property to be narrow in Putnam's sense, it is doubtful that the condition is sufficient. Consider, for instance, the property of being mentally aware of the oldest philosopher. It might well seem that this property should not count as narrow, even though a person's having it would not conceptually imply the existence of any physical objects, or objects other than the person and his own mental states, acts, and experiences. Similarly, the property of knowing that all crows are black and the property of knowing that there are no physical objects might also seem not to be narrow, even though having one of these properties would also not imply the existence of any external objects.¹⁴

Now it is perhaps not entirely clear whether such properties should be counted as narrow or not. After all, each of the properties mentioned would be "narrow" at least in the sense that a person's having the property would not conceptually imply the existence of any objects external to the person's mental experience. But if we are inclined, as I am, to *not* count such properties as narrow, then I think it is because we are assuming that a narrow psychological property must be a property that one can *know* one has solely on the basis of internal evidence.

For instance, one could not know that one knows that all crows are black without knowing that all crows are black. In turn, one could not know that all crows are black without knowing that no crows exist, or that many crows have been observed and all of them have been black, or something of the sort. But it surely seems impossible to know things like this without knowing that the external world exists; and perhaps it is for this reason that the property of knowing that all crows are black does not strike us as being narrow. If so, then a narrow psychological state should be a state that one can know one is in independently of one's having any knowledge of the external world.

With this in mind, then, let us define Putnam's notion of a narrow psychological state as follows:

- (D1) The psychological property of being S is *narrow*₁ if and only if:
- (1) the property of being S is not contingently relational; and
 - (2) necessarily, for any person x, if x is S, then it is possible for x to know that he himself is S without knowing any proposition that conceptually implies either that
 - (a) there exist physical objects, or that
 - (b) there exist contingent objects other than x and x's own mental states, acts, and experiences.

I call the concept defined 'narrow₁' so as to distinguish it from a second concept of narrowness, to be introduced shortly.

5. RETURN TO PUTNAM'S ARGUMENT

If we understand Putnam's Twin Earth argument against meaning-internalism to be stated in terms of narrow₁ psychological states, then I think we must agree that Putnam's argument is successful. So understood, the conclusion of that argument is that the following version of meaning-internalism is false:

- (MI-1) Necessarily, if two persons share all the same *narrow*₁ psychological states, then one of the persons means something by a given word if and only if the other person means the same thing by that word.

I think it is pretty clear that Putnam's Twin Earth case provides a counterexample to (MI-1). Notice that the fact that Oscar and Toscar would not share all the same first person states cannot be used to block this result, since first person states, being contingently relational, are by definition not narrow₁.

But as important as this result is, we should not overestimate its scope. For a slightly weaker version of meaning-internalism, a version that is *compatible* with the consequences of the Twin Earth case, may be obtained simply by adopting a slightly weakened concept of a narrow psychological state. The concept in question would allow certain contingently relational states, such as first person beliefs, to be narrow, provided that the states in question conceptually imply the existence only of particular objects that occur within the agent's immediate mental experience.

Now the concept of an object's being within an agent's immediate mental experience would seem to be the same as Russell's concept of

acquaintance. (Russell, 1912, Chapter 5.) For Russell, we are acquainted precisely with those things that directly present themselves to our minds, things such as ourselves, our own mental acts and states, abstract entities like propositions and universals, and the constituents of sense perception, which Russell called 'sense-data.'

Using Russell's notion of acquaintance, then, let us define the concept of a "narrow₂" psychological property as follows:

- (D2) The psychological property of being S is *narrow*₂ if and only if: *either* the property of being S is *narrow*₁, *or* the following three conditions all hold:
- (1) the property of being S is contingently relational; and
 - (2) necessarily, for any contingent object y, if for all objects x, the existence of y is conceptually implied by the assumption that x is S, then necessarily, no person z is S unless z is acquainted with y; and
 - (3) necessarily, for any person x, if x is S, then it is possible for x to know that he himself is S without knowing any proposition that conceptually implies either that
 - (a) there exist physical objects, or that
 - (b) there exist contingent objects other than x and x's own mental states, acts, and experiences.

The distinctive feature of this definition is the presence of clauses (1) and (2), which allow *narrow*₂ states, unlike *narrow*₁ states, to be contingently relational and to conceptually imply the existence of particular objects of acquaintance. (Clause (3) is necessary in order to rule out states that, while satisfying clauses (1) and (2), nevertheless conceptually imply, for instance, that there are physical objects.)

The concept of a narrow psychological state defined by (D2) provides a second version of meaning-internalism, namely:

- (MI-2) Necessarily, if two persons share all the same *narrow*₂ psychological states, then one of the persons means something by a given word if and only if the other person means the same thing by that word.

For the reasons stated earlier (in Section 2), the Twin Earth argument is ineffective against a form of internalism like this. For again, it is quite implausible to suppose that the protagonists of the example, Oscar and Toscar, would share all their first person cognitive attitudes in common. (Indeed, they would do so only if neither *had* any such attitudes.) Thus the case cannot provide a counterexample to (MI-2), since Oscar and Toscar would not share all their *narrow*₂ states in common.

I think it is clear that if (MI-2) is true, then the basic insight of the received view remains intact. This is because narrow₂ psychological states, even when they are contingently relational, are still *internal* in an important sense. Thus if a person's linguistic understanding is always determined by the person's narrow₂ states, then that understanding must be considered logically independent of the causal, historical, or social relations that the person bears to other persons or objects in the external environment.

The distinction between principles (MI-1) and (MI-2) corresponds to the contrast between two historically important and distinct lines of thought that occur among proponents of the received view. One of these lines of thought originates with Frege and the other originates with Russell. On Frege's view, the cognitive attitudes are all narrow₁, since according to Frege (1892), the contents that essentially characterize such attitudes are complex abstract senses (propositions), all of whose components are also abstract senses. Russell (1912), by contrast, allows the cognitive attitudes to essentially involve both abstract entities (universals) and certain concrete, contingent objects, namely, objects of acquaintance such as selves, mental acts, and sense-data. Thus for Russell, some cognitive attitudes are contingently relational and so are not narrow₁, though all cognitive attitudes are narrow₂. Among proponents of the view that a person's linguistic understanding is determined by that person's cognitive attitudes, a Fregean would thus have to endorse (MI-1), while a Russellian would instead endorse (MI-2).

Given our discussion to this point, then, we may conclude that Putnam's Twin Earth argument refutes *Fregean* internalism, but so far leaves *Russellian* internalism unscathed.

But while the Twin Earth example does not conclusively refute internalism, it certainly leaves the internalist with a serious *problem*. In particular, the internalist needs to provide a plausible explanation of how, in the Twin Earth example, the differences in the way Oscar and Toscar understand the meaning of 'water' can be accounted for by the differences in their first person cognitive attitudes. In the absence of such an explanation, it is open to the externalist to maintain that first person attitudes are simply irrelevant to determining linguistic understanding, and that both (MI-1) and (MI-2) are false.

Providing the relevant kind of explanation is in fact a far from trivial task, since it requires both a theory of meaning for natural kind terms and an account of how such meanings can be grasped and determined by private first person attitudes, even though the meanings in question belong to words in *public* languages. However, I have elsewhere argued in favor of a conceptual framework that I believe is adequate to provide such an explanation (McKinsey, 1986 and 1987). In the remainder of

this paper, I will show how this framework can be applied to the defense of (Russellian) internalism.

Part II. Grasping Public Meanings in Private States

6. SEARLE'S REJOINDER TO PUTNAM

At this point, it will be useful to consider Searle's (1983) attempt to rebut Putnam's Twin Earth argument. For first, Searle's attempted rebuttal is quite similar to the response I gave above that is based on the concept of a narrow₂ state; and second, Searle's rebuttal is inadequate for reasons that clearly bring out some serious problems for any view like mine that might attempt to use the concept of a narrow₂ state to explain how the meanings of natural kind terms are grasped and determined.

Searle accepts (at least for the sake of argument) Putnam's suggestion that a natural kind term such as 'water' might have its meaning specified by an "ostensive definition" such as:

- (8) For any object x , x satisfies 'is water' if and only if x has the same structure as *this stuff*,

where '*this stuff*' is understood as a demonstrative that is used to refer to a particular sample of water. (See Putnam, 1975, pp. 146–150; Searle, 1983, pp. 206–208.) According to Searle's view of indexicals, the use of a demonstrative like 'this stuff' purports to refer to something in the speaker's environment *via* that thing's assumed causal connection with one of the speaker's current perceptual experiences (1983, pp. 218–230). Thus, according to Searle, a semantic rule like (8) resolves into a rule that involves a particular perceptual experience, such as

- (9) For any x , x satisfies 'is water' if and only if x has the same structure as the stuff that is causing this visual experience.

On Searle's hypothesis, Oscar and his counterpart on Twin Earth, Toscar, would grasp different meanings of 'water' via their grasps of the different ostensive rules that each would express by (9). The rules are different, because Oscar's rule concerns one of *Oscar's* visual experiences, while Toscar's rule concerns one of *Toscar's* visual experiences. Moreover, the two rules determine different extensions. For Oscar's visual experience is caused by stuff that is H₂O, while Toscar's visual experience is caused by stuff that is XYZ.

Now in order for Searle's hypothesis concerning the rule (9) to provide a successful rebuttal of Putnam's general rejection of internalism, it

must be possible to use the hypothesis to explain how, in the Twin Earth case, Oscar's and Toscar's understandings of the meaning of 'water' are both different from each other as well as independent of all external objects. But this is not possible. For given the hypothesis, Oscar and Toscar mean different things by 'water' only because they give different *private* meanings to the word: each grasps a different rule that refers to his own visual experience. Given that Oscar and Toscar are grasping such private meanings, then Searle is obviously correct that each person's understanding would be independent of any external objects. Thus *if* Putnam's Twin Earth case were merely an example in which the speakers were grasping different private meanings, then the case could not justify Putnam's denial of internalism.

But this is a big 'if'! For in fact, of course, the protagonists of Putnam's story are grasping different *public* meanings. This is because each person is grasping the meaning that 'water' has in the version of English that he speaks. Thus, Oscar means by 'water' what 'water' means in the English spoken on Earth, while Toscar means by 'water' what 'water' means in the English spoken on Twin Earth—and these meanings are different. So it is open to Putnam to maintain that each speaker's understanding of the public meaning of 'water' in his language must be determined by the speaker's relations to public, external objects, and hence not by any states that are either narrow₁ or narrow₂. Searle's point that in *his* case, each speaker's grasp of a *private* meaning is independent of external objects is a point that, although true, is quite irrelevant to the conclusion of Putnam's argument.

Moreover, I think it is fairly easy to see that, given an assumption that both Searle and Putnam seem to share, Putnam's general denial of internalism really would follow: it would follow from Putnam's example that some meanings can only be grasped by wide, contingently relational mental states that are essentially related to objects outside the head, states that are neither narrow₁ nor narrow₂. The assumption in question is the commonly held view that every cognitive attitude state is essentially characterized by its propositional content. Let us call this assumption 'the proposition theory.' I propose to avoid Putnam's general conclusion by denying the proposition theory. This tactic, however, is not open to Searle, since he explicitly accepts the proposition theory. (See Searle, 1983, Chapter 1.)

7. OBJECTUAL MEANINGS

I have argued elsewhere (McKinsey, 1987) that the Twin Earth case shows that some words have what I call *objectual linguistic meanings*. An "objectual" meaning is a meaning that is given by a semantic rule that essentially involves a particular contingent object, or has such an

object "as a constituent." Like a singular proposition, such a rule can only be expressed by the use of a genuine term such as a demonstrative or proper name, a term whose sole semantic contribution to what is expressed by sentences containing the term is the term's *referent*.

Consider again Searle's proposal that Oscar and Toscar can mean different things by 'water' because they have in mind different rules of the form (14). The rules in question are different simply because they concern numerically distinct (though qualitatively identical) visual experiences. By assumption, Oscar and Toscar associate exactly the same *qualities* with the word 'water.' Clearly then, to distinguish the rules they follow from each other, we need to do as Searle suggests and find different *objects* that the rules might involve. As we've seen, however, the objects suggested by Searle will not do, for being private perceptual events, these objects cannot be constituents of public meanings.

The only alternative that I see is to assume that Oscar and Toscar are following distinct rules that are qualitatively identical but that essentially involve *distinct public objects*. Let 'Q' abbreviate the qualities that are commonly associated with 'water,' such as the qualities of being odorless, wet, thirst-quenching, tasteless, etc. Then one possible hypothesis would be that the rules that Oscar and Toscar follow are of the form:¹⁵

- (10) For any x, x is to satisfy 'is water' if and only if x has the same structure as liquid that is Q on *this planet*.

On my view of demonstratives (unlike Searle's), the rule that Oscar would express by (10) is an objectual rule that would essentially involve the planet Earth. Toscar, on the other hand, would express by (10) a different rule that essentially involves Twin Earth. Notice that what is crucial is *not* the idea that the rules in question are themselves "indexical" in some sense, but rather that the rules can only be expressed by use of *genuine terms* that directly refer to certain objects. Thus, I assume that the rule that Oscar would express by (10) would also be expressed by replacing the demonstrative 'this planet' in (10) with a *proper name* of the planet Earth, or indeed by replacing 'this planet' with any other genuine term that directly refers to Earth. (I do not mean to suggest that we could use (10) to express the rule that we actually do follow in using 'water.' It is in fact implausible to suppose that an understanding of the ordinary meaning of this word requires the ability to think of the planet Earth. I just intend (10) to illustrate the *kind* of rule that governs the use of natural kind terms.)

Now by virtue of what kind of mental state could one be said to grasp or understand such a meaning? Or to rephrase the question: If the meaning of 'water' were given by a rule such as (10), then by virtue of what mental state would it be true, say, that *that* is what Oscar means

by 'water'? One plausible candidate for the relevant kind of mental state would be Oscar's *intention* that tokens of 'water' be subject to the rule in question. If the rule is the one Oscar would express by (10), for example, then perhaps the state by which Oscar grasps this rule could be ascribed by

- (11) Oscar intends that: for any x , x is to be taken as satisfying 'is water' if and only if x has the same structure as liquid that is *Q* on *Earth*.

And now we are close to Putnam's conclusion. Just assume the proposition theory:

- (12) Every cognitive attitude state is essentially characterized by its propositional content,

and

- (13) The propositional content of a cognitive attitude state in which a speaker grasps a semantic rule is that very rule itself.

Then it follows that some meanings are graspable only by contingently relational cognitive attitudes that essentially involve particular external objects. Consider the intention ascribed by (11), for instance. The content of that intention is an objectual rule whose existence conceptually implies the existence of the planet Earth. But if the intention in turn essentially involves the rule in question, then having that intention also conceptually implies the existence of the planet Earth. Thus, it seems that some meanings are graspable only by wide, contingently relational cognitive attitude states that conceptually imply the existence of particular objects outside the head. Such states are of course neither narrow₁ nor narrow₂, and so given this line of reasoning, no form of internalism, whether Fregean or Russellian, seems defensible.

Putnam's general denial of internalism is indeed difficult to avoid. Given the Twin Earth case, we are in my opinion driven to accept this denial, provided that we accept the standard account of the propositional attitudes, according to which the proposition theory is true. And again, since Searle himself believes the proposition theory, there is simply no feature of his views that could allow him to escape Putnam's conclusion.

8. MENTAL ANAPHORA

Our discussion clearly brings out two serious problems for any view like

mine that would try to rebut the Twin Earth argument by appeal to narrow₂ states that essentially involve internal, private objects. First, such a view must be able to explain how it is possible to grasp *public* meanings on the basis of these *private* states. Second, since the relevant meanings are objectual meanings that essentially involve external objects, such a view must be able to explain how these *external* meanings can be grasped by completely *internal* states.

Now although these requirements might at first seem impossible to fulfill, I have elsewhere described a phenomenon that provides a simple solution to both problems (McKinsey, 1986). I call this phenomenon 'mental anaphora.' The most important feature of mental anaphora is the evidence it provides that, contrary to the proposition theory, there are cognitive acts and states that one can be in, even though these cognitive acts and states have *no* propositional contents.

Consider sentences of the following sort:

(14) Oscar wishes he had caught the fish that got away.

What is the propositional content of the wish that (14) ascribes to Oscar? The content cannot be expressed by use of the description 'the fish that got away,' for then the proposition would be, for some time *t*,

(15) Oscar caught at *t* the fish that got away from him at *t*.

But (15) cannot be the content of Oscar's wish, since (15) is contradictory, and Oscar's wish is perfectly consistent. I suggest that we follow Geach (1967) and analyze (14) as

(16) Oscar assumes that just one fish got away, and Oscar wishes it had been the case that he caught *it* (that very fish).

Now suppose Oscar is right that just one fish got away, at time *t*, say. Let us call this fish 'Bubbles.' Then it is clear what proposition Oscar would be wishing true. It is the proposition that is true at a possible world *w* just in case in *w*, Oscar catches Bubbles at *t*. But now suppose that Oscar is wrong. Perhaps what Oscar thought was a fish on the end of his line was really just an old boot or an underwater branch. Then it seems quite impossible to specify the proposition that Oscar wishes had been the case.

Moreover, or so I have argued (McKinsey, 1986), the best account of the semantic behavior of the anaphoric pronoun in contexts like (16) implies that if Oscar's assumption is false, then his wish has no proposition as content. On this account, the second occurrence of 'it' in (16) has the meaning of what Gareth Evans (1977) called an "E-type

pronoun." An E-type pronoun is an indexical genuine term whose reference is fixed by the definite description that is "recoverable" from the clause governed by the pronoun's quantifier antecedent. The idea is that in (16) we utilize a form of pronominal anaphora to ascribe what in effect is a semantically analogous form of *mental* anaphora that occurs in Oscar's mind.

If Oscar's assumption that just one fish got away is false, then the pronoun 'it' as it occurs in 'he caught *it*' in (16) has no referent. And since 'it' is meant here as a genuine term, the sentence 'he caught *it*' would thus express no proposition. Oscar's wish, which involves a mental act of reference analogous to the use of a genuine term whose reference is fixed by a false descriptive assumption, would therefore also have no propositional content. Yet in these same circumstances, (14) and (16) would remain *true*, so that Oscar would still have a perfectly good wish, namely, the wish ascribed by both (14) and (16). Thus it is possible to have a particular wish, without there being any proposition that one is wishing true. And a similar result follows for all cognitive acts and states that, like belief and unlike knowledge, can fail to have true propositional contents. Thus the proposition theory, which says that every cognitive attitude act or state is essentially characterized by its propositional content, is false.¹⁶

The phenomenon of mental anaphora can be used to explain how it is possible for a speaker to grasp an objectual meaning in a cognitive attitude state whose own existence is logically independent of that meaning. Consider again the sort of objectual rule expressed by (10). Oscar might grasp this rule by virtue of his being in the mental state ascribed by

- (17) Oscar assumes that there is just one planet that he himself inhabits, and Oscar intends that: for any *x*, *x* is to be taken as satisfying 'is water' if and only if *x* has the same structure as liquid that is *Q* on *that planet*.

If we continue to assume that Oscar is an inhabitant of Earth, then the rule that is the content of Oscar's intention is an objectual rule that essentially involves the planet Earth. For this is the kind of rule that is expressed by the sentence in the scope of 'intends that' in (17), given that 'that planet' is a genuine term that refers to the planet Oscar inhabits. Nevertheless, I think it is clear that (17) could remain true, and thus Oscar would remain in the same mental state that (17) ascribes to him, in a possible world in which Oscar's assumption is false and he inhabits no planet at all. (Perhaps in such a world, Oscar is just a brain in a vat, floating in the empty void.) Thus the mental state by which Oscar grasps a public meaning of 'water' is a mental state that does not essentially

have that meaning as its content, and it is a mental state whose existence does not presuppose the existence of any objects outside the head.

The most widely advertised consequence of the Twin Earth example is supposed to be that "meanings just ain't in the head" (Putnam, 1975, p. 144). This seems to mean that some words' meanings are not graspable or determined by mental states that essentially involve only things that are in the head. But the phenomenon of mental anaphora shows, I think, that this conclusion does not follow. What *does* follow is the important *semantic* result that some words have objectual linguistic meanings. But given mental anaphora, we can see how such essentially external meanings can be split off from the mental states that grasp them, so that the states themselves remain logically independent of the external world.¹⁷

The mental state ascribed to Oscar by (17) is a complex cognitive attitude consisting of a first person descriptive assumption and a semantic intention that is based on this assumption. Since Oscar is the only contingent object whose existence is conceptually implied by a person's being in this state, it is clear that the state is narrow₂. Oscar's twin Toscar would of course be in an exactly similar narrow₂ state, but since Toscar's first person assumption would be about himself, the rule grasped in Toscar's semantic intention would be a different objectual rule that essentially involves the planet that Toscar inhabits, Twin Earth.

Thus the fact that Oscar and Toscar would be in different first person cognitive states explains how they could grasp distinct public meanings of 'water' even though their internal states are qualitatively and conceptually indistinguishable.

Moreover, the hypothesis that something like this is in fact going on in the Twin Earth case is, I think, independently plausible. As Russell (1912, pp. 55–56) pointed out, one's thoughts about particular things in the external world seem to be typically based upon descriptive assumptions that involve direct reference to objects of acquaintance, including one's self and one's mental states and experiences. If this is so, then it should not be surprising that our grasp of public meanings that essentially involve particular external objects should also be based in part upon the kind of private descriptive assumptions on the basis of which we typically think of those very same objects. For obviously, if we are to think of such meanings, we must first be able to think of the particular objects that the meanings essentially involve.¹⁸

* * *

In this paper, I have tried to defend an internalist point of view regarding the determination of linguistic understanding. If this defense is correct, then we also have no reason to reject an internalist point of view regarding the determination of the meanings of words in public languages. For

it remains plausible to suppose that a word's meaning in a language is determined by the linguistic understanding of the particular speakers of the language, together perhaps with other narrow states of the speakers, such as their communicative intentions.¹⁹

On the other hand, we should not understate the significance of Putnam's argument, since it has forced us to abandon some of the most natural and widely held versions of internalism, resulting in a weakened as well as a sharpened formulation of the traditional point of view. Still, the basic tenet of internalism remains intact. The source of meaning is found in the internal psychological states of individual speakers. There is thus no reason yet to believe that adequate theories of meaning and linguistic understanding will ever have to appeal to any kinds of external causal, historical, or social relations: we are, that is, not yet driven to accept any kind of "causal theory" of either meaning or linguistic understanding.²⁰

Wayne State University
Detroit, Michigan

NOTES

¹ Eloquent modern expressions of this point of view can be found in the work of H. P. Grice (1957, 1968, 1969) and David Lewis (1969, 1975).

² Strictly, this follows only on the assumption that 'water' is not an indexical term whose meaning remains constant while its extension varies from context to context (planet to planet?). Some philosophers have in fact suggested that 'water' and other natural kind terms are indexicals. (See, for instance, Almog (1981) and Campbell (1982).) But in my opinion, this suggestion is very implausible. For a convincing criticism of the hypothesis that natural kind terms are indexicals, see Burge (1982), pp. 103-107.

³ The logical properties of such sentences were first discovered and exhaustively explored by Castaneda (1966, 1967, 1968). See also Perry (1977, 1979).

⁴ The fact is, then, that there are always *two* distinct properties that first person belief sentences like (1) and (2) ascribe to their subjects. This is not a characteristic that is unique to such sentences, for it is a characteristic possessed by other simple relation-sentences that contain a reflexive pronoun. Consider, for instance, the sentence (i) 'Ollie shaves himself.' On the one hand, (i) attributes to Ollie the property of being a self-shaver, a property possessed by all and only those who shave themselves. But on the other hand, (i) also attributes to Ollie the quite different property of being an Ollie-shaver, a property possessed by all and only those who shave Ollie. That (i) attributes this second property to Ollie is made clear by the fact that (i) expresses the same proposition as (ii) 'Ollie shaves Ollie.' The situation regarding first person belief sentences like (1) and (2) is thus a special case of a quite unproblematic fact about all reflexive relation-sentences.

⁵ Since the view of first person belief just endorsed is similar to the views of Lewis (1979) and Chisholm (1981), I should perhaps briefly indicate how my view differs from theirs. I have followed Lewis and Chisholm in that, central to both my view and theirs is the concept of a person's self-ascribing a property. (Thus, my property H is just the property

that any person x has if and only if x self-ascribes the property of being a hero.) However, my view differs crucially from theirs in its implication that each first person belief essentially involves, or is essentially *about*, the person who has the belief. Because it has this implication, my view accounts for the difference of the beliefs that are held when each of two persons believes, say, that he himself is a hero. Failing to have this implication, the views of Chisholm and Lewis apparently cannot explain this difference in belief.

Indeed, Lewis seems to explicitly deny the difference, and seems to claim that two persons, each of whom believes that he himself is a hero, would have precisely the *same* belief (Lewis, 1979, p. 143). But it seems to me that this claim is just false. For it seems to me that it is the *semantic* properties of beliefs that determine the beliefs' sameness or difference. Thus, if two persons have the same belief, then the beliefs in question necessarily have the same *truth-value*, are necessarily *about* the same objects, and necessarily concern the same *propositions*. Since the first person beliefs of two persons do not necessarily share these semantic properties, even when they involve self-ascription of the same properties, the beliefs in question are always counted as different beliefs.

The general view of first person belief that I would endorse is very similar to the view proposed by Perry (1979). Like Perry, I would emphasize that a first person belief always has as its content a certain singular proposition about the believer. Thus the belief ascribed by (1) 'Ollie believes that he himself is a hero,' has as its content the proposition that Ollie is a hero. But also like Perry, I would add that believing a singular proposition about oneself is not all there is to having a first person belief: one must also believe the proposition *from a first person perspective*. We thus might describe the property which on my view is ascribed by (1) by saying that it is the property of believing the proposition that Ollie is a hero from a first person perspective.

On the other hand, my view of first person belief is embedded in a general view of the cognitive attitudes that is apparently quite different from Perry's. For reasons that I discuss elsewhere (McKinsey, 1986) and that I outline in Section 8, I reject the general principle, apparently endorsed by Perry, that the proposition believed is always an essential characteristic of the belief in question. I will discuss first person belief further in note 17.

⁶ This argument, as well as other relevant considerations, may be found in Pollock (1982), pp. 14–16.

⁷ This problem with (3) was brought to my attention by Barbara Humphries.

⁸ The significance of such properties was brought home to me by Lawrence Powers.

⁹ My skepticism regarding (7) was encouraged by Jay Rosenberg.

¹⁰ There is still another conception of a "narrow psychological state" that has commonly occurred in the recent literature concerning Twin Earth problems. On this conception, a "narrow" psychological state is, by definition, a psychological state that is supervenient upon neurophysiological states, so that molecular duplicates must of necessity share all the same narrow psychological states. (See, for instance, Block (1986), p. 620 and Fodor (1987), pp. 30–32.) I have not discussed this conception in the text, primarily because, being explicitly "physicalist," the conception seems irrelevant to the issue that Putnam initially raised about the traditional view of how meaning is determined. Thus, on this physicalist conception, the Twin Earth case shows that linguistic understanding is not always determined by psychological states that are supervenient on brain states. But it is far from clear how this conclusion is supposed to be related to the views of such philosophers as Frege, Russell, and Grice about the psychological determination of meaning.

On at least some concepts of supervenience it follows from the physicalist conception of narrowness that a person's being in any given narrow psychological state would (by definition of 'narrow') *entail the existence of certain physical things*, namely, the person's own brain and its physical states. (This would follow, for instance, on Kim's (1984) concept of strong supervenience.) But surely, such a conception of narrowness is entirely contrary in spirit to the Cartesian concept of narrowness that Putnam initially introduced and in

terms of which he raised a problem for the traditional view of meaning. In this connection, note also that the physicalist conception of narrowness is subject to counterexamples of the same kind as those that undermine the proposal that a psychological state is narrow iff it is not contingently relational. Consider, for instance, the property of having a thought that is caused by an event in one's own brain. Surely, this is a psychological property that molecular duplicates would share in common, and so the property would be narrow according to the physicalist conception. But of course since having the property logically implies the existence of certain physical objects (i.e., brains), the property would certainly not count as narrow in Putnam's original sense.

The physicalist conception might of course be useful in some contexts, such as Fodor's (1987) discussion of psychological explanation. But for the reasons just given, it is of no use to me in this paper. My goal here is to motivate and provide clear explications of narrowness that can in turn be used to state clear versions of the traditional view of how meaning is determined, the view that Putnam takes himself to have refuted. If, as I contend, one of the versions of the traditional view that I explicate turns out to be unaffected by any considerations surrounding Twin Earth examples, then I take it that this will be an important result, regardless of any verbal disputes that might arise over how we ought to use the word 'narrow'.

¹¹ Le Pore and Loewer (1986, pp. 610–611) distinguish between what is metaphysically possible and what is epistemically possible relative to one's knowledge of one's own mental states. The distinction, of course, is ultimately due to Kripke (1972).

¹² Here I assume that each step of such a correct deduction would follow from previous lines by a self-evident inference rule of some adequate system of natural deduction.

¹³ My student David Shier reminded me of this fact.

¹⁴ The property of knowing that all crows are black was suggested to me by Lawrence Powers, and the property of knowing that there are no physical objects was suggested to me by Gerald Powers.

¹⁵ For a reason I discuss in (1987), p. 11, a rule of the form (10) cannot be correct when construed as a *definition* of 'water.' If *this planet* = Earth, then the definition would be wrong, since it is possible for stuff that is XYZ to be liquid that is Q on Earth; but such stuff would nevertheless not be water. My proposal in (1987) is that the semantic rule for 'water' does not *define* 'water' but instead *fixes the property* that 'water' is used to predicate (1987, pp. 11–16). However, this complication need not concern us here, and (10), even though it is not correct, can at least serve to illustrate the notion of an objectual rule.

¹⁶ A further consequence of mental anaphora is that the cognitive attitude verbs such as 'wishes,' 'thinks,' and 'believes' do not express mental relations that hold between persons and propositions. It is thus incumbent upon me to provide an alternative account of the meaning and logical form of cognitive attitude ascriptions. In my initial (1986) paper on mental anaphora I made a brief (and inadequate) suggestion, and I can do little better here. Following Sellars, I would suggest that the cognitive attitude verbs are predicate-forming operators that, when attached to a sentence, form a predicate that we use to classify a mental state in semantic terms, terms that involve an analogy with the semantic properties of the sentence in question. (To believe that p is to believe in a that-p manner, where this does not require the existence of the proposition that p; we might call this an "adverbial theory," as opposed to a relational theory, of belief. See, for instance, Sellars (1969).) The important fact brought out by mental anaphora is that the analogy must be based on semantic features of the operated-upon sentence other than the proposition it expresses. Instead, I would say, the analogy is based upon those semantic features of the sentence that *determine* which proposition it expresses.

One plausible initial hypothesis would be that the analogy is based upon the relevant sentence's context-independent linguistic meaning. (The meaning would have to be context-

independent, it seems, since no belief could be fully characterized semantically by a context-dependent meaning that determines many distinct propositions relative to different contexts.) But while I think that this hypothesis is on the right track, it cannot be completely correct. For sometimes, belief ascriptions are made with belief-predicates formed from sentences that *have no* context-independent meaning. An especially poignant kind of example would be sentences containing anaphoric pronouns, such as the second conjunct of (16) in the text. Again, I cannot give an adequate account here, though I hope to do so in a forthcoming work now in progress.

¹⁷ Similarly, the phenomenon of mental anaphora allows us to see how it is possible for us to believe singular propositions about external objects, even though those beliefs are logically independent of the singular propositions and external objects in question. This in turn makes it possible for the fundamental semantic properties of our cognitive attitude states to always be quite *internal*, in the sense of always being at least narrow₂, even though the contents of these attitudes may be singular propositions that essentially involve external objects. It also makes it possible for us to have privileged, non-empirical knowledge of the fundamental semantic properties of our own cognitive attitude states (since these properties are always narrow₂), even when the objects of the beliefs are singular propositions to which we have *no* privileged access. (See McKinsey (1990).)

Sometimes, though, beliefs whose contents are singular propositions are also essentially characterized by those propositions. An example of this that is especially important for the present discussion would be first person beliefs. If what I said earlier in the text is correct, a first person belief is always essentially characterized by its being *about* the believer in question. The existence of the belief, therefore, guarantees the existence of the subject of the singular proposition believed, and hence guarantees the existence of this proposition (assuming that the proposition involves the simple ascription of a one-place predicate). For instance, in every possible world in which Ollie believes that he himself is a hero, the proposition that Ollie is a hero exists and is the content of this belief.

So *de se* beliefs are *exceptions* to my general rule that typically, beliefs in singular propositions are logically independent of those propositions. Given my thesis that all beliefs are fundamentally characterized by properties that are narrow₂, all beliefs that are like *de se* beliefs in being essentially characterized by the singular propositions believed must also be beliefs in propositions whose contingent constituents are all *objects of acquaintance* for the believer. The resulting view, then, is very closely related to that of Russell (1912). But I take it that my view has a significant advantage over Russell's in that my view allows us to believe singular propositions whose contingent constituents are *not* objects of acquaintance. In such cases, of course, the proposition believed is not an essential feature of the belief, and so this advantage over Russell is purchased at the price of giving up the proposition theory as a *general* account of belief.

¹⁸ For a recent and very persuasive argument that our thought about external objects fundamentally depends upon our ability to make first person references, see Pollock (1988), pp. 188–189. A similar point of view is also expressed by Chisholm (1981), Chapter 4.

¹⁹ In the manner suggested by Grice (1968) and Lewis (1969, 1975).

²⁰ A short version of Part I of this paper was presented to the Central Division of the American Philosophical Association under the title "The Internal Basis of Meaning" (New Orleans, April, 1990). Most of the material in Part II was presented to the Central Division of the American Philosophical Association as part of my comments on Michael Devitt's paper, "Meanings Just Ain't in the Head" (Cincinnati, April, 1988). I am grateful to Richard Feldman, Barbara Humphries, Lawrence Lombard, Gerald Powers, Lawrence Powers, Jay Rosenberg, and my student David Shier for useful comments and discussions concerning these matters.

REFERENCES

- Almog, Joseph: 1981, "Dthis and Dthat: Indexicality Goes Beyond That," *Philosophical Studies* 39, 347-381.
- Block, Ned: 1986, "Advertisement for a Semantics for Psychology," *Midwest Studies in Philosophy* 10, 615-678.
- Burge, Tyler: 1982, "Other Bodies," in A. Woodfield (ed.), *Thought and Object: Essays on Intentionality*, Clarendon Press, Oxford, 97-120.
- Campbell, John: 1982, "Extension and Psychic State: Twin Earth Revisited," *Philosophical Studies* 42, 67-89.
- Castaneda, Hector-Neri: 1966, "'He': A Study in the Logic of Self-Consciousness," *Ratio* 8, 130-157.
- Castaneda, Hector-Neri: 1967, "Indicators and Quasi-Indicators," *American Philosophical Quarterly* 5, 85-100.
- Castaneda, Hector-Neri: 1968, "On the Logic of Attributions of Self-Knowledge to Others," *Journal of Philosophy* 65, 439-456.
- Chisholm, Roderick: 1981, *The First Person: An Essay on Reference and Intentionality*, University of Minnesota Press, Minneapolis.
- Donnellan, Keith: 1970, "Proper Names and Definite Descriptions," *Synthese* 21, 335-358.
- Donnellan, Keith: 1974, "Speaking of Nothing," *Philosophical Review* 83, 3-31.
- Evans, Gareth: 1977, "Pronouns, Quantifiers, and Relative Clauses (I)," *Canadian Journal of Philosophy* 7, 467-536.
- Fodor, Jerry A.: 1987, *Psychosemantics*, MIT Press, Cambridge, Mass.
- Frege, Gottlob: 1892, "On Sense and Reference," in *Translations From the Writings of Gottlob Frege*, Basil Blackwell, Oxford, 1966.
- Geach, P. T.: 1967, "Intentional Identity," *Journal of Philosophy* 64, 627-632.
- Grice, H. P.: 1957, "Meaning," *Philosophical Review* 66, 377-388.
- Grice, H. P.: 1968, "Utterer's Meaning, Sentence-Meaning, and Word-Meaning," *Foundation of Language* 4, 1-18.
- Grice, H. P.: 1969, "Utterer's Meaning and Intentions," *Philosophical Review* 78, 147-177.
- Kaplan, David: 1968, "Quantifying In," *Synthese* 19, 178-214.
- Kaplan, David: 1979, "On the Logic of Demonstratives," in P. French, T. Uehling, and H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, University of Minnesota Press, Minneapolis, 401-412.
- Kim, Jaegwon: 1984, "Concepts of Supervenience," *Philosophy and Phenomenological Research* 45, 153-176.
- Kripke, Saul: 1972, "Naming and Necessity," in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, D. Reidel, Dordrecht, 253-355.
- Le Pore, Ernest and Loewer, Barry: 1986, "Solipsistic Semantics," *Midwest Studies in Philosophy* 10, 595-614.
- Lewis, David: 1969, *Convention*, Harvard University Press, Cambridge, Mass.
- Lewis, David: 1975, "Languages and Language," in K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science* 7, University of Minnesota Press, Minneapolis, 3-35.
- Lewis, David: 1979, "Attitudes De Dicto and De Se," in David Lewis, *Philosophical Papers*, Volume I, Oxford University Press, Oxford, 1983.
- McKinsey, Michael: 1986, "Mental Anaphora," *Synthese* 66, 159-175.
- McKinsey, Michael: 1987, "Apriorism in the Philosophy of Language," *Philosophical Studies* 52, 1-32.
- McKinsey, Michael: 1991, "Anti-Individualism and Privileged Access," *Analysis* 51, 9-16.

- Perry, John: 1977, "Frege on Demonstratives," *Philosophical Review* 86, 474-497.
- Perry, John: 1979, "The Problems of the Essential Indexical," *Nous* 13, 3-21.
- Pollock, John: 1982, *Language and Thought*, Princeton University Press, Princeton, NJ.
- Pollock, John: 1988, "My Brother, The Machine," *Nous* 22, 173-212.
- Putnam, Hilary: 1970, "Is Semantics Possible?" in S. Schwartz (ed.), *Naming, Necessity, and Natural Kinds*, Cornell University Press, Ithaca, 1977.
- Putnam, Hilary: 1975, "The Meaning of 'Meaning'" in K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science* 7, University of Minnesota Press, Minneapolis, 131-193.
- Russell, Bertrand: 1912, *The Problems of Philosophy*, Oxford University Press, New York, 1959.
- Searle, John R.: 1983, *Intentionality*, Cambridge University Press, Cambridge.
- Sellars, Wilfrid: 1969, "Metaphysics and the Concept of a Person," in Karel Lambert (ed.), *The Logical Way of Doing Things*, Yale University Press, New Haven.