PHYSICAL EXTERNALISM AND SOCIAL EXTERNALISM: ARE THEY REALLY COMPATIBLE?

JEELOO LIU
SUNY, GENESEO

ABSTRACT: In this paper I examine the foundations of physical externalism and social externalism and argue that these foundations are incompatible. Physical externalism is based on a direct reference theory of natural-kind terms, while social externalism is based on a description theory of natural-kind terms. Thus, physical externalism and social externalism are incompatible just in the same way that the direct reference theory of proper names is incompatible with the description theory of proper names. My argument will proceed as follows. In Section One, I shall explain what the two theses say and spell out my suspicion. In Section Two, I shall take a look at the initial setups for physical externalism and social externalism by examining Putnam's and Burge's original arguments. Finally in Section Three, I shall explain that the real incompatibility comes to lie in the different assumptions on which the two theories are based. I will present some thought experiments to highlight this incompatibility.

Putnam and Burge have been viewed as launching a joint attack on individualism, the view that the content of one's psychological state is determined by *what is in the head*. Putnam argues that meanings are not in the head, while Burge argues that beliefs are not in the head either, and both have come up with convincing arguments against individualism. It is generally conceived that Putnam's view is a version of physical externalism, which argues that factors in the physical environment play a role in determining the meanings of natural-kind terms. Burge, on the other hand, is regarded as following up Putnam's argument to bring in factors in the social environment

for the determination of belief. Burge's view has been commonly referred to as "social externalism." The general consensus in the field is that physical externalism and social externalism are compatible views. Furthermore, both Putnam and Burge seem to endorse each other's position for the most part. In this paper, however, I shall argue against this general view to show that the two theories are deep-down incompatible.

In my presentation of such a view, it has often been pointed out to me that either Putnam later revised his theory to accommodate Burge's intuition, or that Burge in some of his works appeals to Putnam's intuition. However, my contention here is not that Putnam's overall philosophy and Burge's overall philosophy are incompatible with each other. My contention is rather that physical externalism is based on a direct reference theory of natural-kind terms, while social externalism is based on a description theory of natural-kind terms. And thus, physical externalism and social externalism are incompatible just in the same way that the direct reference theory of proper names is incompatible with the description theory of proper names.

My argument will proceed as follows. In Section One, I shall first explain what the two theses say and spell out my suspicion. In Section Two, I shall take a look at the initial set-ups for physical externalism and social externalism by examining Putnam's and Burge's original arguments. I will explain how Burge's initial project is totally different from that of Putnam's, and then I will show how his view, at times, contradicts Putnam's conclusion. Finally in Section Three, I shall explain that the real incompatibility comes to lie in the different assumptions on which the two theories are based: Physical externalism is based on an *essentialist's* account of natural-kind terms; social externalism is based on a *description theorist's* account of general kind terms.

I. VARIETIES OF EXTERNALISM: PHYSICAL AND SOCIAL

To begin with, *externalism* is not the obvious claim that the contents of our thoughts, desires, etc. are affected by what goes on in the external environment. *They are*. Not even individualists would deny this claim. Our beliefs *are* about physical objects with which we interact; our expressible thoughts *are* related to concepts that we share with our linguistic community. These claims are uncontroversial. However, individualists argue that such a causal/conceptual relevance is already captured *internally*, so that we only need to look internally to establish our mental taxonomy. Under the individualist's interpretation, most of the differences in the physical environment are already reflected in the individual's brain states through the individual's interaction with the physical environment; most of the differences in the social environment are already reflected in the individual's conceptual scheme through the individual's acquisition of her linguistic knowledge. Externalism, on the other hand, is the claim that the individuation of our thoughts is affected by what goes on in the external environment,

even when differences in the external environment are not reflected inter-nally. Since most of the differences in the external environment are already reflected in the individual's brain states, the kind of scenario externalists can suggest has to be farfetched. To prove externalism, they need to have ex-amples that can capture the possibility of having some environmental changes affecting the individuation of thoughts, even when the individual herself is completely ignorant of such changes. Externalists also need to provide a justification or an explanation for their individuating mental states (or men-tal contents) differently, when there is no epistemic difference on the individual's part. That is, they need to give us a good story.

Physical externalists can give us two stories. On the one hand, they can claim that an individual's mental content is affected by differences in the physical environment, in the way that her mental content can change *simply as a result of her being situated in a different physical environment*. In this case, there is a *direct tracking* of mental content to physical objects,[1] such that the objects are presented to the individual *immediately*. The "immediacy of objects" means that reference is not mediated through other members of the social/linguistic society. I shall call this view *direct physical external-ism*. For example, let us suppose that Oscar put a glass of water by his bed before going to bed. Upon his waking, he saw the glass of water and thought to himself: "It is so nice to be able to drink water before even getting up." According to this version of physical externalism, Oscar's present thought is *not* the same type as Oscar's thought "It is so nice to be able to drink water before even getting up" if he were to be transported (without his awareness) to Twin Earth in his sleep. Even though Oscar's thought (interpreted internally) did not change; even though his concept of *water* did not change, his thought on Twin Earth is viewed as a different type of thought simply because it is now related to a different object. *Direct physical externalism* bases the indi-viduation of mental states on the individuation of physical objects. One can-not have a thought about a physical object if one is not directly related to the object(-type) itself.

On the other hand, physical externalism can also try to establish the con-nection between mental content and physical objects by tracing the causal/historical chain (in the spirit of Kripke). Oscar does not have to have direct contact with the object in question in order to have his thought related to the object; his thought can be related to that object as long as he got his concept from others who have had direct contact with the object. There is, under such a version of physical externalism, a *parasitic causal/historical* connection between Oscar's thought and its proper object. I shall call this view *historical physical externalism*. Under *historical physical externalism*, Oscar's thought on Twin Earth could initially be typed as the same thought as on Earth, until he later establishes contact with the linguistic community on Twin Earth and starts to use their concept of *water*.

Social externalists can claim that an individual's mental content is affected by changes in her social environment when the societal conceptual framework differs in any respect. Such a difference in the societal conceptual framework brings about different individuation of the individual's mental contents, even when the individual herself does not have any self-aware change of her understanding or perspective. For example, Oscar could be thinking that "Lemon is good with Gin & Tonic" and this thought would be a different type of thought when the societal usage includes what we call 'lime' under their term 'lemon', even when Oscar himself is thinking only about that yellow acidic fruit. That is to say, Oscar could be in the same physical environment, his thought could be about the same kind of object, and he could be having the same understanding of the term. His thought would change simply as the result of changes outside of his internal states (nonintentionally specified), because his thoughts are related to the world *indirectly* through his intellectual/conceptual framework. And such a framework is determined *socially*.

Under *direct* physical externalism, there is a direct relation between the individual's mind and the world. The individual's mental content changes (becomes a different type) as the *direct* result of the individual's being situated in a different physical environment. The individual does not need to rely on a causal chain established through history, just as she does not need to rely on social conventions, to be related to objects in the physical environment. In other words, under *direct* physical externalism the individual's mind is "self-contained with respect to the doings of other speakers." Under historical physical externalism and under social externalism, on the other hand, such a *self-containedness* is being denied.[2] I suspect that the incompatibility between physical externalism and social externalism will manifest itself in their giving different answers to the following question: is the individual's mind directly tracking down physical objects, or is it tracking down physical objects through the medium of others' reference in the community? I shall argue that physical externalism, at least as initially extended from Putnam's Twin Earth Argument, should be construed as a form of *direct* physical externalism. Social externalism, on the other hand, places the individual's mind in a socially constructed intellectual/conceptual framework and assumes an *indirect* connection between the mind and the world. I shall turn to my explication of the two theories next.

II. PUTNAM'S THEORY VS. BURGE'S THEORY

In "The Meaning of 'Meaning'," Putnam develops an *indexical* theory of meaning.[3] An indexical term is one that has a fixed linguistic meaning while its extension changes depending on the context of usage. The fixed linguistic meaning Putnam (1975, 231) assigns to 'water' is this:

(SL) (For every world W)(For every x in W)(x is water $\equiv$ x bears *same$_L$* to the entity referred to as 'this' *in the actual world W$_1$*)

The *same$_L$* (same liquid) relation is a certain equivalence relation. This specification in (SL) has an indexical nature, in that it includes an ostensive designator "this." The "this" in the actual world—our Earth, as we know, refers to H_2O. However, Putnam argues that this specification has at the same time a *rigid* nature, in the sense that once it picks out the relevant object in the actual world, the extension no longer shifts from one possible world to another. That is to say, 'water' is indexical to our world and its reference is thereby fixed for all possible worlds. Since the liquid referred to as 'water' on Twin Earth does not bear the *same$_L$* relation to the water in the actual world, it is not water, and cannot be said to be the extension of 'water'. It is in view of this that Putnam says that the term 'water' does not have the same meaning in the two worlds, even though what Earthians and Twin Earthians have in mind (or the stereotypes they associate with the term) are the same.

Putnam's *indexical* theory of meaning is established for natural-kind terms. When it comes to natural-kind terms, Putnam thinks that neither the individual's concepts, nor the socially defined intensions (the Fregean *sense*), could determine their meaning. The meaning of such a term is partially determined by the nature of the physical environment that surrounds us. In the original dubbing (supposedly there was one) a natural-kind term was coined to pick out a particular set of things that have in common a certain *stuff*. This term thereupon designates this particular set of things. Reference-fixing is not dependent on descriptions, even though a "descriptive definition" was presumably given initially.[4] This description is supposed to provide the *criteria* in accordance with which the right set of things will always be picked out. The speakers are thereby given a means by which they could tell whether a thing belongs to that set or not. However, neither could the original description, nor could any later revised description, really *define* the meaning of a natural-kind term. Depending on our knowledge of the object, the descriptions of the proper criteria for determining the scope of things might vary. But *the stuff* is always there, and its essential qualities would not change whether we discover them or not. This being the case, a natural-kind term may have different descriptive definitions in different societies, though it has but one extension. With further studies, there may eventually come a description that correctly *fits* the nature of the natural kind in question. Therefore, as far as a natural-kind term is concerned, what its extension is should be determined, not by each linguistic society's definition of the term, but by the natural boundaries that exist among the nature of things. It is thus the scientists' job, not that of the linguists', to find out what each natural kind is. As Putnam puts it: "What *really* distinguishes the classes we count as natural kinds is itself a matter of (high level and very abstract) scientific investigation and not just meaning analysis." (Putnam 1970, 141)

Under this indexical view of meaning, meaning is external not only to an individual's *head* but also to society's *heads*. Social/linguistic conventions cannot completely determine *meaning*, because the physical nature of the things being referred to is also part of *meaning*. The most important aspect of Putnam's Twin Earth argument is not to show that meaning is not *individualistic*, but to show that meaning is not in the *heads*, or as Putnam puts it, the *collective mental states*, of all the members of the linguistic society. In the Twin Earth thought experiment, we see that the two societies may apply exactly the same criteria ("is transparent, is odorless, is drinkable, is what is in the lakes, rivers, etc.") in determining what their term 'water' refers to. But the extensions of these two terms are different; hence, their *meanings* are different. Since what the two societies think about their liquid is the same, meaning is in no way determined by what the society thinks about the extension, or what they associate with the term 'water'. Thus, Putnam's externalistic theory of meaning really comes out to be a *nonmentalistic theory of meaning*. Meaning does not lie *in the head*, because it lies *in the physical nature of things*. In this respect, Putnam's externalistic theory of meaning can be properly described as "physical externalism."

This externalistic theme for Putnam was originally set up as a theory of meaning, not as a theory of mental content. As a theory of mental content, Putnam's position seems to remain *individualistic*.[5] In Putnam's view, two people may have the same concept even when the extensions differ. For instance, Putnam argues, if he thinks "I have a headache," and his *Doppelgänger* thinks "I have a headache" too, "the same word, 'I', has two different extensions in two different idiolects; but it does not follow that the concept I have of myself is in any way different from the concept my *Doppelgänger* has of himself." (Putnam 1978, 131) What is more, different words can be associated with the same concept in one person if that person does not know the difference in the extensions. Using the example of the concepts *elm* and *beech*, Putnam says that if a speaker does not know how elm and beech differ, then "the conceptual content associated with the words 'elm' and 'beech' is practically the same." (1975, 226) The concept or the conceptual content, in Putnam's usage, may not capture either the *intension* or the *extension* of the term. It embodies, more precisely, how *the individual* understands the intension, or what she/he associates with the extension. It is in this sense that concepts are said to be "in the head." Furthermore, suppose that the correct extensions of 'beech' and 'elm' are reversed on Twin Earth, and Putnam and his *Doppelgänger* are both unable to tell the difference, Putnam argues, "It is absurd to think his psychological state is one bit different from mine." (1975, 226) From this we see that psychological states for Putnam are basically *individualistic*.

Even though Burge used the same Twin Earth example in "Other Bodies," he changed the original scenario by adding one more specification of the

differences between Earth and Twin Earth: The scientists have discovered the relevant facts about water (twater) and thus the social meanings of the two terms are different in the two societies. It is significant that Burge stipulates that there are differences between the scientific communities in the two scenarios. This suggests that Burge does not think of extension as a component of meaning. If he did, then, like Putnam, he might have concluded that 'water' has different meanings in the two scenarios just on the basis of the difference in extensions.

As we saw earlier, Putnam's Twin Earth thought experiment is a case of contrasting the physical environment with the social environment. Burge, however, is interested in the contrast between the social environment and the individual's thought. His inquiry using this argument is this: if the social concept of the term 'water' has changed, would the water-related thoughts of those individuals *who have not yet obtained such knowledge* be affected? His answer is naturally "yes." According to Burge, what specifies the content of a mental state is the *obliquely occurring* expression in the that-clause of a belief report.[6] To demonstrate that one's mental content is not determined completely by what goes on in one's brain, Burge wants to find a possible situation where an expression occurring obliquely in the that-clause remains the same from the individual's perspective, while the meaning of the expression changes in two different social settings. Thus, he came up with a thought experiment with the *arthritis* case.[7]

What is creative in Burge's thought experiment is his focus on the legitimacy of the usage of words. Burge is not saying that *our* English word 'arthritis' adopts a new meaning in the counterfactual situation; he is on the contrary arguing that *that word* is no longer a word in English. He suggests that we coin a new word 'tharthritis' to represent it in English. This move is by now a commonly accepted move in understanding a thought experiment of this kind, but its revolutionary spirit has not been fully appreciated. What Burge is saying is that even while constructing a thought experiment, we cannot forget that the words we use in our discourse are governed by the conventional rules of the language of our discourse, and in this case, English. We cannot simply conceive of a thought experiment in which *our words* adopted some new meanings; instead, we must always say that those words with different social meanings are simply *not* our words in English.

Extending from Putnam's claim that meaning has to be studied nonindividualistically, Burge claims further that beliefs have to be studied nonindividualistically as well. The meaning or concept in one's belief varies with differences in social connotations, thus the content of one's belief is determined by what goes on in the social environment. When one's belief contains a misunderstood or misapplied notion, Burge argues that we should still give the belief a *literal* ascription, unless the misunderstanding is too deviant for the subject to be ascribed a belief as such. His *fortnight* example

demonstrates the point that literal interpretation is prima facie the preferred interpretation, as long as the speaker is a normal, competent member of the linguistic community.[8] In most cases, "communal conventions about the meaning of a speaker's words tend to override what a speaker mistakenly associates with his words in determining what he says and even, sometimes, believes." (Burge 1978, 134–35)

Burge's attack on individualism is not built as a theory of meaning, but as a theory of mental content. He wants to emphasize that whatever expressions mean in the social environment, that will be what they mean in the content of the individual's belief. When there is a gap between the social meaning and the believer's understanding of the word, it is the social meaning, and not what the individual understands, that determines what the individual really believes. Burge's conclusion is thus: *beliefs are not in the head of the individual*. Putnam calls his own theory "semantic externalism," since it is a theory about meaning. In contrast, Burge's theory should be called "mental externalism," since it is a theory about mental content.

With regard to the issue of mental content, Putnam and Burge have their initial disagreement.[9] Putnam says that $Oscar_1$ and $Oscar_2$ have the same *psychological states in the narrow sense*. By "psychological states in the narrow sense" Putnam means leaving out the consideration of extension. Contrary to Putnam's claim, Burge argues that even on this level, the Earthian's and the Twin Earthian's beliefs involving the notion *water* cannot be said to be the same. So Burge is actually using Putnam's example to argue against the very assumption that Putnam holds about concepts, namely, the assumption that concepts are in the individual's mind. While what Putnam claims is that extension is not determined by the content of psychological states, *since content is totally in the head*, what Burge claims is that extension is not in the individual's head, *and content is not either*. Concepts for Burge are thus *nonindividualistic* entities. Since Putnam's *concepts* are to be interpreted as individualistic psychological entities, we should conclude that Burge's philosophical adversaries include even Putnam.

With regard to the issue of meaning, the two views are also in disagreement. (Later I will show how it is in this difference that we can discover the incompatibility between physical externalism and social externalism.) Putnam's original thought experiment is put forth to dispute the view that "the collective mental states" of society can determine meaning, whereas Burge altered the original thought experiment to prove that meaning *is* determined by the collective mental states of society. Furthermore, in Putnam's picture, as long as water is in my physical environment, I am having *direct access* to this kind of stuff. My concept of water is about water, and I don't need to depend on others' usage for *access to the referent* of 'water.' But in Burge's picture, the *mediation* of others' usage seems to play a crucial role in determining what kind of concepts I have and what kind of referents can be picked out by my concepts.[10]

It may be argued that Putnam's theory does not rule out the role of social environment since he believes in *the division of linguistic labor*. I think it does have a sociolinguistic aspect, but as Bilgrami points out, Burge's social externalism project goes much farther than Putnam's *division of linguistic labor*.[11] Furthermore, social externalism and the thesis of *the division of linguistic labor* make different claims on different issues. Social externalism is based on a general sociolinguistic theory of meaning, while the *division of linguistic labor* is a specific thesis on some special terms that involve sophisticated scientific knowledge. For instance, Putnam says that words like 'chair' do not exhibit this division of labor, and that 'water' "did not exhibit it at all before the rise of chemistry." (Putnam 1975, 228) That is to say, his Twin Earth example as set at the time frame of 1750 does *not* involve the division of linguistic labor. Earthians and Twin Earthians share the same social conventional meanings of the term 'water' (in Burge's terms, they are *in the same social environment*), and yet Putnam claims that the meaning of 'water' differs in the two worlds. This shows that *meaning* for Putnam is simply *not* a mere sociolinguistic issue.

For Putnam, the division of linguistic labor "rests upon and presupposes the division of *non*linguistic labor," and by "nonlinguistic labor" Putnam means the labor of scientists. The role of scientists is to investigate the physical environment. According to Putnam's indexical theory, extension is determined by the physical environment, not by what scientists say—since they could be mistaken. Thus, even scientists as experts cannot define the extension (and thus the meaning) of a natural-kind term for us. Scientists' nonlinguistic labor is the foundation of linguists' linguistic labor. And since these discoveries may not be complete yet, there may always be some *indeterminacy* in the result of our linguists' linguistic labor.[12] The application of the word would change with each new discovery, but the whole point of Putnam's indexical theory is that the extension of the term 'water' does not change along with the experts' discovery. Extension is simply not some Fregean function of our experts' knowledge. Finally, Putnam uses this hypothesis to demonstrate that "everyone to whom gold is important for any reason has to *acquire* the word 'gold'; but he does not have to acquire the *method of recognizing* whether something is gold." (Putnam 1975, 227–228) The division of linguistic labor comes merely as the laymen's relying on the experts' providing superficial properties for recognition, while the experts' real contribution is to investigate the essential properties. Even when the experts could provide criteria or methods of recognition, they do not provide meaning. As Stephen Schwartz points out, "The extension of the term is part of the meaning, but criteria specifiable in terms of superficial properties are not." (1978, 81) In conclusion, the hypothesis of the division of linguistic labor is simply not a sociolinguistic theory of *meaning*.

Perhaps we can wrap up the present section in this way: Putnam's physical externalism is an *anti-individualistic* and *anti-sociolinguistic* theory of meaning, but he holds an *individualistic* theory concerning concepts and psychological states *in the narrow sense*. Burge's social externalism is an *anti-individualistic* and *sociolinguistic* theory of meaning, and he also holds an *anti-individualistic* theory concerning concepts and all psychological states. If one accepts Putnam's physical externalism, then one would have to deny Burge's assumption that meaning is simply "linguistic meaning." If one accepts Burge's social externalism, then one would have to deny Putnam's assumption that mental concepts are individualistic entities. They may be holding the same "anti-individualism" banner on the surface, but basically they are dealing with different issues and have different concerns.

III. THE INCOMPATIBILITY BETWEEN PHYSICAL EXTERNALISM AND SOCIAL EXTERNALISM

So far I have only argued that the two views draw different conclusions both on how meaning is determined and on how mental content is individuated. In this section, I will show how the two views are ultimately incompatible. To make my point, let me first introduce a brief thought experiment that I shall call the *expert error* example.[13]

Suppose that at the time (1750?) when our scientists made the discovery about the nature of water, *the chemical components of water on Earth were actually XYZ*, but somehow these scientists had simply made a mistake, and the mistake was *never* uncovered. Since social conventions about the meaning of natural-kind terms were directed by the general consensus among scientists, our notion *water* included this information that water was H_2O. In this counterfactual situation, Putnam would say that 'water' in English never did *mean* H_2O. Burge, on the other hand, would have to say that 'water' still meant H_2O and we were not talking about 'water' in this thought experiment.

This *expert error* example can first of all demonstrate how the two views are different theories of meaning. For Putnam, what determines whether something qualifies as water is not what we say about water, but what essence the thing *initially dubbed as 'water'* actually has. There is no guarantee that the meanings of all natural-kind terms have been completely exhausted or correctly captured by our scientific discoveries. Meaning thus cannot be fixed onto our *present* linguistic conventions, and even now we may not know the meaning of all the natural-kind terms in our language. In this way Putnam seems to take meaning to be a variable, the value of which is determined by the nature of the physical environment. Our ignorance of meaning is a direct result of our ignorance of the physical environment. Given that such an ignorance is possible, the *expert error*, however unlikely, is conceivable. If in our world (i.e., Earth) there were such an expert error at the initial stage of dubbing,

then all the descriptive definitions given in our present dictionary, including that of 'H_2O', would not be the meaning of 'water'.

With Burge, the scenario would be given a different interpretation. As he did with his *water* example or his *arthritis* example, Burge would ask us to consider this case in a *metalinguistic* way. He would say that in our construal of this thought experiment, we are imagining the liquid on Earth to be different from what we call 'water' in our language, and thus we need to coin a new word for this imaginary liquid. *That* XYZ liquid in our thought experiment is simply not what *our water term* would pick out; therefore, we cannot call it 'water'. I think Burge would concede that we may still be very ignorant of the nature of physical objects, but I don't think he would take this to imply our ignorance of *meaning*. That is to say, even if we don't give the thought experiment a metalinguistic rendition, Burge would agree with Putnam that in this scenario water *is* XYZ, but he would not say that the term 'water' *means* XYZ. Conventional linguistic meaning for Burge is a matter of linguistic conventions, and we (as a society) know what linguistic conventions our society has. Linguistic conventions are of course subject to changes under new discoveries of physical sciences. However, meaning is not changed *before* a new discovery is made known to the society. Even after a new scientific discovery is made, meaning can be changed only *as a result of* the change of social conventions. Suppose there were this expert error in the initial dubbing, and our scientists found out about the error. After the news is spread to the majority of our society, the lexicographers would have to revise the definition of 'water' in our dictionaries, and *then* 'water' in English would mean XYZ. Revision of meaning occurs as the result of such a continuous process.[14]

The *expert error* example can also demonstrate how a physical externalist would disagree with a social externalist on the individuation of mental content. Physical externalism, as a theory of mental content, could be broadly construed as the thesis that *the content of one's mental state changes as the result of differences in the physical environment, even when such differences are not reflected in one's internal states* (*namely, what is in one's head*).[15] The thesis of social externalism, on the other hand, could be put this way: *the content of one's mental state changes as the result of differences in the social environment, even when such differences are not reflected in one's internal states* (*nonintentionally specified*). Thus in the present *expert error* case, a physical externalist would take the individual's mental content to remain the same, since the individual is situated in the same physical environment. A social externalist, on the other hand, would clearly decide that the individual's mental content cannot be the same since in the actual world, it involves the notion *water*, while in the hypothetical scenario, it does not.

The *expert error* example presents a case where the social environment *misrepresents* the physical environment, and it is in such a case that the two theories may be in conflict. Burge argues that the physical environment could

be *mediated* through the social environment, and both environments affect the individuation of mental content. Most externalists (physical and social alike) seem to embrace this assumption about *mediation,* taking it for granted that the social environment does reflect the physical environment in the correct way. However, when there is a mismatch, as presented in our *expert error* example, one has to decide whether mental content is determined *primarily* in accordance with the physical environment, or with the social environment.

One thing I should stress is that the incompatibility emerges only when *natural-kind terms* are involved in the mental content.[16] It has been pointed out that the foundation for physical externalism is Putnam's *scientific essentialism* concerning natural-kind terms.[17] The foundation for social externalism, I shall argue, is a kind of *Fregean description theory* of general kind terms. Under the essentialist's view, natural objects are marked by their natural boundaries. A scientist investigating these natural boundaries aim to capture the *hidden* microstructure of natural objects. The natural boundaries already exist among different kinds of objects, and the scientist's job is simply to match natural-kind terms with these natural kinds.[18] On the other hand, under the description theorist's view, it is the *description* that determines the boundary of a natural kind or the extension of a natural-kind term. That is to say, objects fall into the extension of a natural-kind term *in virtue of satisfying a certain set of necessary and sufficient conditions*. Change the description of these conditions (albeit nonarbitrarily), and you can get a larger or a smaller extension for the term. Therefore, essentialism and the description theory (concerning natural-kind terms) are two competing views.[19] Next I will explicate the connection between physical externalism and scientific essentialism.

Putnam argues that the taxonomy of natural kinds is based on the nature of the stuff, and the natural boundaries among kinds are marked by the physical make-up or the essential qualities those objects share in common. The essential properties or essences could be interpreted as the genetic make-up in the case of living things, and atomic or molecular properties in the case of natural substances.[20] We have already explained that Putnam thinks it is the scientific experts' job to investigate those essences. With regard to the issue of natural-kind terms, that is, with regard to how these terms designate the kind (or the essence of the kind), Putnam thinks that there is only one correct theory: the *direct reference* theory. Putnam takes natural-kind terms to be terms that pick out "natural" kinds; that is, kinds that are individuated by nature. The reference of a natural-kind term is directly fixed through some kind of *ostensive* definition, not through any *descriptive* definition. For Putnam, natural-kind terms are determined by the natural boundaries between individual objects and kinds in the context of the *initial* usage of these terms.[21] In other words, when the words are first coined to designate a certain kind of object, there are differences between the nature of this kind of object and that of another kind. Even if the initial users of these words are not aware of the

differences, the words themselves are "tags" for the different natures. Since Putnam assumes that there are such natural boundaries, natural-kind terms for him would have different meanings as long as they were originally fixed onto different natural kinds. The difference is established even when we fail to give enough information to identify the boundaries between the two kinds. As Kripke puts it, "The original concept . . . is: *that kind of thing*, where the kind can be identified by paradigmatic instances. It is not something picked out by any qualitative dictionary definition." (Kripke 1972, 122) Two things may have the same qualitative properties and yet differ in kind; they can also be of the same kind and yet exhibit different qualitative properties. In Kripke's words, it is only the *substance* that matters.[22] Reference for natural-kind terms is thus determined, not through descriptions of any qualitative property, but though a direct fixing onto the essences.

Naturally, the substances or the essential properties are not just the superficial properties known to laymen. However, they are not necessarily known to our experts either. If we think that our experts have already given us full knowledge of the essences of natural kinds, and a thing belongs to a particular natural kind as long as it meets the *experts' descriptions of the essences*, then we are misconstruing Putnam's and Kripke's direct reference theory. Reference is *direct* exactly because it is not *mediated* by any description, whether it is a description of superficial properties or a description of essential properties. As Abbott explains, the relation between natural-kind terms and essence under Putnam-Kripke's theory is this: "[E]xactly because the real essence was unknown, it could not possibly be semantically associated with the kind term. It is as if, in the semantics of these expressions, in place of the denotation determining properties we have simply a variable standing for an unknown quality." (Abbott 1989, 272) This unknown quality is what we call the "underlying structure" or "essence." Putnam does not take these structures to be unknowable, but his theory leaves open the possibility that *our* science and *our* language do not capture them yet. Since the correct usage of natural-kind terms relies on scientific discoveries, there is indeed a need for constant revisions in the usage of natural-kind terms. In other words, the *nonlinguistic* labor will take some time before it is incorporated into the division of linguistic labor. But the change of scientific theories does not bring about different referents. Putnam's and Kripke's direct reference theory holds that there is a "referential stability" (borrowing Arthur Fine's terminology) when such terms undergo scientific revisions. Reference is determined through the initial usage of the term, and the revision is just to better match the boundary that is there *in nature*.

Physical externalism, as a natural extension from Putnam's scientific essentialism and his direct reference theory, should be *direct* physical externalism. If the *essence* we are directly related to does not consist in superficial macroproperties, then our thoughts about natural kinds are not determined by

what we (individually or collectively) perceive or know about the superficial macroproperties. But under scientific essentialism, there is a possibility that the essence is not even defined by the *presently known* microproperties. Any property of water, deemed essential or not, could turn out *not* to be true of water. Putnam (1975, 161) calls this possibility an "epistemic possibility" (as opposed to "epistemic necessity"). Take "Water is H_2O" for example. If it is not epistemically necessary that water is H_2O, then "Water is H_2O" is not analytic. The two terms 'water' and 'H_2O' are co-referential in our world, but 'H_2O' does not give the meaning or provide a conceptual analysis of our notion *water*. Putnam (1975, 244) says, "The reason we *don't* use 'cat' as synonymous with a description is surely that we know enough about cats to know that they have a hidden structure, and it is good scientific methodology to use the name to refer rigidly to the things that possess that hidden structure, and not to whatever happens to satisfy some description." Putnam holds that there is a *co-reference* relationship between 'water' and 'H_2O' or between 'cat' and 'a feline animal', etc. But this co-reference relationship is simply not the same as synonymy. Therefore, even descriptions of essential properties do not *give the meaning* of natural-kind terms.

Let us briefly go back to the case of Oscar's being transported to Twin Earth in his sleep. If his thought upon waking up (that is, before he has any linguistic contact with Twin-Earthians) is a different thought from *what he would be thinking if he were still on Earth*, then the change is brought about by his direct rapport with (the essence of) the object itself. There is no *mediation* through *the linguistic community*—whether it is represented by what the majority thinks about the macroproperties, or by what the experts say about the essence of the stuff.

Description theories of natural-kind terms, on the other hand, treat such terms as *connotative*. A *connotative* term is one that is associated with a description that defines a set of necessary and sufficient conditions.[23] This set of descriptions constitutes the *sense* of the term, and the *sense* determines the extension for the term. This view assumes that the extension of a natural-kind term is *mediated through* descriptions, and there is no direct reference between the object and the term. I will now argue that the very foundation of social externalism lies in this *connotative* or *description* theory of general kind terms.

If Burge had employed the direct reference theory of natural-kind terms, then he would not have needed a separate discussion on how one's mental content could be affected by differences in one's social environment, when the content itself involves a natural-kind term. Differences in the physical environment alone are sufficient to bring about differences in the mental content. Burge's strategy in his arguments clearly shows that he did not employ such a direct reference theory. Furthermore, Burge takes statements like "Water is H_2O" and "Cat is a feline animal" to be statements that give us

synonymies in the same way that the statement "Bachelors are unmarried males" does. In other words, in each of the three statements, the term (the subject) and the description (the predicate) *have the same meaning*. To avoid the problem Quine sees in defining "having the same meaning"; namely, the problem of meaning's lacking an ontological status, Burge argues that two terms are synonyms if the most competent speakers would use them *interchangeably*. In his words, "Synonymies are grounded in practice." (Burge 1986a, 701) Even though Burge denies that these statements can be true simply in virtue of having this synonymy relationship, he does not deny that the predicate description *gives the meaning* of the subject term. For example, the dictionary definition of 'water' in our present English is this: "colorless liquid compound of oxygen and hydrogen (H_2O); liquid consisting chiefly of this and found in seas and rivers, in rain, and in human tears, etc." This description gives us what the term 'water' *means* in our present English. Thus a statement like "Water is H_2O" is a "meaning-giving" statement in that it gives us two synonyms. This conclusion is clearly just the opposite of Putnam's conclusion as explicated earlier.

Burge's examples demonstrate that the meaning of all terms, natural-kind terms included, changes along with changes in society's conventional usage. For example, what is shown in the way Burge sets up his *arthritis* example is that such a change of conventional usage affects not only the subject's mental content, but also the taxonomy of the kind 'arthritis'. In the counterfactual situation, it is not just that the meaning of 'arthritis' has changed, the extension has also changed as a result of the change of the term's definition.[24] According to Burge, the way we use 'arthritis' in our society and the counterfactual usage of 'arthritis' would pick out different diseases—they have different *extensions*: "For counterpart expressions in the content clauses that are actually and counterfactually ascribable are *not even extensionally equivalent*." (my italics) (Burge 1979, 79) If this is so, then it is our conventional usage of 'arthritis' that determines the extension of 'arthritis'. In other words, *it is* our conventionally assigned qualitative descriptions that give the meaning of the term 'arthritis', and thereby define the scope of its kind. This again supports my interpretation of his theory as the connotative view on natural-kind terms.

It may be disputable whether 'arthritis' counts as a natural-kind,[25] and maybe Burge's thought experiment could be taken to involve only nonnatural-kind terms. However, Burge himself does not think so. He says, "We could have used an artifact term, an ordinary natural-kind term, a color adjective, a social role term, a term for a historical style, an abstract noun, . . . or any of various other sorts of words." (Burge 1979, 79) This quote clearly shows that with any natural-kind term, Burge could have constructed a similar thought experiment in which a change of meaning brings about a change of extension. In "Wherein is Language Social?" Burge also says, "Individuation of an individual's concepts

or translational meaning may depend on the activity of others on whom the individual is dependent for acquisition of *an access to the referents*. If the others by acting differently *had put one in touch with different referents,* compatibly with one's minimum explicational abilities, one would have had different concepts or translational meaning." (my italics) (1989, 187) In other words, individuals do not have direct access to the referents of these terms; reference is mediated through others' usage. Putnam's theory is called "the direct reference theory"; in contrast, Burge's theory could be called "the *mediated* reference theory," or simply, "the *indirect* reference theory." I think that this mediated reference theory is simply a form of the description theory, because it is the descriptions that mediate reference. Under this theory, a different description (i.e., a different social/linguistic meaning) *could* get the individual related to a different referent. There is thus no *referential stability* as seen in Putnam's theory. Furthermore, under Burge's view, descriptions not only *fix the reference*, they also *give the meaning* of natural-kind terms. Even though these descriptions are mostly based on scientific discoveries, it is the descriptions alone, not the so-far-undiscovered essence, that give the meaning of these natural-kind terms.

The main difference between the direct reference theory and the description theory of natural-kind terms is whether the reference is *mediated* through some descriptions. In other words, direct reference theory holds reference to be *direct* in the sense that there is no description (or the Fregean *sense*) to serve as the mediating link between words and objects. Burge has on numerous occasions stated that his theory is simply an extension of Putnam's physical externalism, since his arguments merely have the physical environment *mediated through* the social environment. But it is exactly this kind of *mediation* that separates social externalism from direct physical externalism. Burge's social externalism paints the picture that we *as individuals* are not directly related to the physical environment; our words refer to objects through the mediation of other people's activities. I argue that this picture is contrary to the claim of direct physical externalism.

As stated earlier, the incompatibility between physical externalism and social externalism emerges only in the case of natural-kind terms. I shall further demonstrate this incompatibility by constructing two thought experiments involving natural-kind terms in the spirit of Burge, since he says that all his thought experiments can be applied to natural-kind terms as well. The first thought experiment will proceed in the same manner as Burge's *arthritis* example. As a first step, imagine a person, Jenny, who has a piece of jade for years. The piece of jade Jenny has is actually a piece of green jadeite (a silicate of sodium and aluminum). From observing this piece of jade and other samples, Jenny forms the belief that jade has a high degree of hardness and firmness and is mostly green. One day Jenny decides to buy another piece of jade, so she goes to a jade shop. When the shop owner presents Jenny with

a piece of yellow jade (that is actually a piece of nephrite; a silicate of calcium and magnesium), Jenny claims that it is not jade and the shop owner is trying to deceive her. In our present usage, the word 'jade' includes both jadeite and nephrite. So Jenny's belief is false and the shop owner is being honest. Now as a second step, imagine Jenny or Jenny's physically identical counterpart in the same situation up to the moment when Jenny expresses her disbelief. The only difference is that in this counterfactual scenario, the societal usage is different and the word 'jade' includes only jadeite. Thus Jenny's belief turns out to be true and the shop owner is trying to con her into buying something that is not what that society would call 'jade'. According to Burge, if we compare Jenny's belief in the first and the counterfactual scenarios up to the moment when she expresses her disbelief, we should conclude that in the two scenarios Jenny has different mental contents involving the notion *jade* in the oblique context. The reason is that in the counterfactual situation, the same word form 'j-a-d-e' no longer expresses the same notion as our present English notion *jade*. It should properly be replaced by something like 'T-jade' to mark the fact that it is no longer the same word. According to the theory of meaning underlying this thought experiment, in the first situation the societal usage relates the word 'jade' to a larger kind that includes jadeite and nephrite, while in the second situation, the societal usage relate 'jade' to a smaller kind. As Burge puts it, the expression 'j-a-d-e' picks out different extensions in the two scenarios because of the different societal meanings. This is exactly the view Putnam sets out to refute: *intensions determine extensions*.

The second thought experiment will proceed in the same manner as Burge's *sofa* example. As a first step, let us imagine a person, David, who learns the word 'rat' through ostensive instruction. David has not observed many examples of rat, but from the few examples he has seen, he notices the mechanical behavioral patterns rats typically exemplify. David thus formed the hypothesis that rats are not animals but roborats. David's hypothesis will certainly be falsified once he proceeds to verify it. As a second step, imagine a counterfactual situation in which there were no rats-as-animals; there were only rats-as-roborats on our Earth. So in this scenario, David's hypothesis will turn out to be true and his theory will no longer be nonstandard. But since in our present English usage, the word 'rat' refers to the kind of *animals* we see in subway tunnels, in the field, or even in our own kitchen, Burge would say that David's thought in the second scenario no longer involves the notion *rat*. We should probably coin a new word such as 'T-rat' to express David's hypothesis. And this hypothesis will certainly be verified in this new scenario. This thought experiment demonstrates that words like 'rat', 'cat', 'animal', etc. are defined by our present linguistic conventions. Once they *mean* something different, Burge would just say that we must treat the word as having the same *word form* but expressing a different notion. Under Putnam's theory, on the other hand, *if it were true that* our rats turned out to be roborats,

then the word 'rat' should simply *mean* roborats and its extension would be these roborats.

To sum up, we should go back to the main thesis of externalism: the agent's external environment is relevant to the determination of the content of his/her intentional mental states. I think the key question is *how* the physical or the social environment becomes relevant. The relevance of the physical environment is established through the *intentionality* of the individual's mental content. On the other hand, the relevance of the social environment is established through the individual's conceptual scheme underlying her mental content. I can even construct two arguments for each side against the other side. For example, the physical externalist's argument could go as follows:

(1) The individual's intentional mental states are *about* the physical environment and are *caused by* objects in the physical environment.

(2) One cannot separate mental content from what the mental state is *about*.

(3) This interrelation is primarily established on the individual level: if an individual's mental tokens are about water and not *twater*, or are caused by water and not *twater*, then the individual's intentional mental states are directly related to water.

(4) There is thus no need for social *mediation* in these cases.

In a similar vein, the social externalist's argument could be stated as follows:

(1) The individual as well as society at large is interacting with the world *conceptually*.

(2) The societal conceptual scheme determines the concepts in the agent's mental content; there are no *individualistic* concepts.

(3) Any mental content that involves a concept would have to be affected by the social conventions governing this concept.

(4) Therefore, the social mediation between the individual's mental content and the physical world is required.

IV. CONCLUSION

The debate between the direct reference theory of proper names and the description theory of proper names is well-known. And now we see that this debate also extends to natural-kind terms. In this paper I have argued that physical externalism is based on a direct reference theory of natural-kind terms, while social externalism is based on a description theory of natural-kind terms. And thus physical externalism and social externalism are incompatible just in the same way that the direct reference theory of proper names is incompatible with the description theory of proper names.

It seems that neither Putnam nor Burge saw the incompatibility in their theories. Putnam (1973, 132) says, "[A] better semantic theory must encompass

both [the contribution of society and the contribution of the real world]." I do not contest this claim. It is true that both physical externalism and social externalism play a significant role in challenging individualism. There would be no problem if physical externalism is built on natural-kind terms only, while social externalism is built on nonnatural-kind terms only. We can even call it "a division of externalistic labor." But Putnam basically thinks that what he says about the *indexicality* of natural-kind terms can be applied to other kind terms as well. He even uses 'pencil' as an example, and concludes that "It follows that 'pencil' is not *synonymous* with any description — not even loosely synonymous with a *loose* description. When we use the word 'pencil', we intend to refer to whatever has the same *nature* as the normal example of the local pencils in the actual world. 'Pencil' is just as *indexical* as 'water' or 'gold'." (Putnam 1975, 243) Burge, on the other hand, intends to use the same meaning analysis on artifact terms and natural-kind terms alike. We simply don't see Putnam and Burge themselves separating their external-istic labor. Stephen Schwartz (1978, 87) gives a very pertinent suggestion: "The correct approach to ordinary language . . . is to combine both Putnam's insights and the traditional approach. It is clear that Putnam is correct about natural-kind terms, his error is in extending his analysis to nominal-kind terms. On the other hand, followers of the traditional approach are correct about nominal-kind terms, and their error has been in attempting to extend their analysis to natural-kind terms." I would push the issue further, and ask whether anyone could hold both physical externalism and social externalism when it comes to the meaning of natural-kind terms, or the individuation of mental contents involving natural-kind terms.

Finally, the bottom line of the whole dispute seems to be this: is the mind related to the essential elements in the world *directly*, or is its relation to the objects *mediated though* the conceptual scheme of the mind? In this paper I suggested some finer distinctions could be made in regard to the relevance of external environments to one's mental content. One can be a physical exter-nalist by holding two kinds of views: the first view holds that one's mental content is established through a direct reference (or 'indexicality' in Putnam's usage) relationship; the second view holds that one's mental content is estab-lished through a Kripkean causal/historical relationship with the object. The second relationship would presuppose the first one, but not the reverse. Social externalism would be compatible with the second form of physical external-ism, but not with the first (*direct*) form. A finer distinction could even be sug-gested for social externalism: one could be a social externalist by holding that one's mental content is established through a direct "conceptual fit" relation-ship, or one could hold that the individual's mental content is established through a causal/historical social relationship with the object. Burge's theory seems to be of the second kind. A lot more can be done to distinguish between varieties of externalism. However, I shall not endeavor to do so in this paper.

ENDNOTES

[1]The term 'direct tracking' is from McCulloch (1995, 168). He points out that what is being tracked is the "essences" of physical objects. This significance of this claim will be discussed later in this paper.

[2]McCulloch (1995, 180; 182, note 5) says that Burge's "Individualism and the Mental" denies that the mind is self-contained with respect to the doings of other speakers.

[3]Putnam is using 'indexical' in a way different from the standard usage of the term. It may be argued that Putnam does not really hold an *indexical* theory of meaning because he is using the term differently. However, Putnam himself points out that *indexicality* is the essence of his theory of natural-kind terms. I thus think that calling his theory an *indexical* theory is not doing him an injustice.

[4]According to Putnam (1975, 229–30), this is a description that "typically consists of one or more markers together with a *stereotype*—a standardized description of features of the kind that are typical, or 'normal', or at any rate stereotypical."

[5]John McDowell calls Putnam's view an "isolationist conception of the mind." As McDowell (1992, 315) argues, "[T]he 'isolationist' conception of language that Putnam objects to is all of a piece with a similarly 'isolationist' conception of the mind—at least of the mind as it is in itself. And Putnam's attack on the 'isolationist' conception of language leaves the counterpart conception of the mind unquestioned."

[6]That is to say, for any term in the content clause, one cannot substitute in its place a co-extensive term with a sure guarantee that the whole sentence's truth value would remain intact. Burge specifies that the content of one's mental state is *provided by* the that-clause in the mentalistic ascription.

[7]The thought experiment goes in several steps as follows:

Step 1: Conceive an arthritis patient, S, who has had arthritis in his ankles and wrists for years. S was informed of the diagnosis years ago, and has since formed the belief that "I have had arthritis in my ankles and wrists for years." Based on this information it is legitimate for us to report that:

    (P) S believes that he has had arthritis for years.

"That he has had arthritis for years" provides the content of S's belief in this case.

Step 2: Feeling some discomfort in his thigh, S thought that his arthritis has spread to his thigh. We further report this:

    (Q) S believes that his arthritis has spread to his thigh.

Step 3: In our social convention the word 'arthritis' only refers to a chronic disease that is marked by inflammation and stiffening of the joints. So S's arthritis could not have spread to his thigh. (Q) reports a false belief on S's part. However, S is not aware of the mistake in his understanding of 'arthritis'.

Step 4: Now under the supposition that S's physiological history, behavioral history, physical build-up and epistemic perspectives all remain unchanged, let us imagine a counterfactual situation where the social convention does allow the word 'arthritis' to apply to various rheumatoid ailments, making the belief reported in (Q) a true belief.

Step 5: Run the counterfactual situation parallel to the original situation from step 1. Examine the that-clause in (P), we'll see that since the word 'arthritis' now has a different meaning, (P) is saying something different. That is to say, in this counterfactual situation our attribution, using 'arthritis' in its new meaning, *provides* a different content from what we originally attributed. This difference is brought about by the variation in the individual's social environment alone, since the individual aspect is presumably fixed in both situations.

For further details of this argument, see Burge 1979a.

[8]Here is Burge's example:

> (L)   For years I believed that a fortnight was ten days, not fourteen, though of course I never believed that fourteen days were ten days.

In this case the misunderstood notion was *fortnight*. The speaker meant by it "a period of ten days." Thus if he believed that Bill would be back by a fortnight, he believed that Bill would be back by ten days. This seems to be a reasonable interpretation. But Burge resisted all attempts to interpret the speaker's belief by taking into account what the speaker meant. To him, 'fortnight' *means* "a period of fourteen days," and nothing else. When the speaker's belief content incorporates this notion, it should simply be interpreted according to its conventional meaning. Thus in his interpretation, the speaker did believe that Bill would be back by a fortnight (whatever the word 'fortnight' means conventionally). Burge claims that one's holding beliefs involving notions that one does not fully understand is a common fact of life. Our literal belief attribution will serve to expose this incomplete understanding.

[9]Putnam later seems to accept Burge's intuition on the difference that exists between the mental contents of Oscar[1] and Oscar[2]. But then I don't see how his original argument could proceed.

[10]For instance, Burge (1989, 187) says, "First, in many cases we must, on cognitive grounds, defer to others in the explication of our words. Second, *the individuation of our concepts and meanings is sometimes dependent on the activity of others from whom we learn our words and on whom we depend for access to the referents of our words*." (italics mine)

[11]I came to the same conclusion independently of Bilgrami, but I found that I agree with most of what he says concerning these two forms of externalism. For a detailed explanation of his argument, see Bilgrami 1994, 23–24.

[12]See Abbott's analysis of Putnam-Kripke's theory (Abbott, 1989).

[13]I would like to thank Christopher Gauker for suggesting this name.

[14]For a detailed discussion on meaning revisions, see Poncinie, 1985.

[15]Strictly speaking, such a version of physical externalism is not found in Putnam's original argument, since he is talking about psychological state in the narrow sense and he argues that Earthians and Twin-Earthians *are* in the same psychological states. But this version has been extended from his original argument and he doesn't seem to object to it.

[16]A natural-kind term is a term that designates a class of natural objects, such as 'lemon', 'grass', and 'tiger', etc. To say that this class of objects forms a natural kind, is to claim that the taxonomy is based on some common underlying structure of these objects.

[17]According to Robert Hanna, both Kripke and Putnam hold the doctrine of scientific essentialism, which is glossed by Kripke as such: "In general, science attempts, by investigating basic structural traits, to find the nature, and thus the essence (in the philosophical sense) of the kind." See Hanna, 499.

[18]For example, William Goosens (1977) argues that natural-kind terms lack logically necessary and sufficient conditions, and properties associated with the term cannot determine the extension of the term.

[19]A nonnatural-kind term, on the other hand, picks out either an arbitrary set of things, or an arbitrary subset of a natural kind, by some *stipulative* conditions alone. For instance, 'contract', 'bachelor', 'sofa', and 'chair' are nonnatural-kind terms, and they seem to be determined through a set of descriptions the social conventions assign to the objects. If our usage of 'sofa' includes 'chair', for example, then this new notion picks out a new kind (both sofas and chairs fall into this kind) and it is then not the same notion as *sofa* used in present English. There doesn't seem to be any insurmountable obstacle to our society's changing the definition to include a larger or a smaller nonnatural kind. This is because artifacts, customs, or social practices are simply made by humans and are categorized according to human interests. We can simply list properties we deem to be relevant to our purposes, and our descriptions of these properties determine the extension of these nonnatural-kind terms. Thus, for nonnatural-kind terms, there doesn't seem to be two competing views, or at least the essentialist view does not seem to apply.

[20]I am here using Abbott's interpretation. See Abbott, 271.

[21]The emphasis on the *initial* usage is intentional here. In some of Putnam's examples (such as his examples of 'cats'), he argues that if there is a natural mutation in the essential properties of cats (for example, they were originally robocats and were later replaced by the organic cats that we see around us), the meaning of 'cats' in our English should *mean* the initial samples (namely, robocats).

[22]Kripke (1972, 128–29) says, "[I]f this substance can take another form—such as the polywater allegedly discovered in the Soviet Union, with very different identifying marks from that of what we now call water—it is a form of water because it is the same substance, even though it doesn't have the appearances by which we originally identified water."

[23]I am using the term 'connotative' as Kripke characterizes it. See Kripke 1972, 127.

[24]Burge (1979a, 78) says, "The person might have had the same physical history and nonintentional mental phenomena while the word 'arthritis' was *conventionally applied*, and defined to apply, to various rheumatoid ailments, including the one in the person's thigh, as well as to arthritis." (italics mine)

[25]Presently the term is defined through the symptom (the inflammation of the joints) of the disease. If it is discovered that arthritis is caused by a certain virus, then whenever the inflammation is caused by this virus, it is a case of arthritis. 'Arthritis' would then be a natural-kind term.

BIBLIOGRAPHY

Abbott, Barbara. 1989. "Nondescriptionality and Natural Kind Terms." *Linguistics and Philosophy* 12: 269–91.

Bilgrami, Akeel. 1994. *Belief & Meaning*. Oxford: Blackwell Publishers.

Burge, Tyler. 1972. "Truth and Mass Terms." *The Journal of Philosophy* 69: 263–83.

———. 1975. "Mass Terms, Count Nouns, and Change." *Synthese* 31: 459–78.

———. 1977. "Belief De Re." *Journal of Philosophy* 74: 338–62.

———. 1978. "Belief and Synonymy." *Journal of Philosophy* 75, no. 3: 119–38.

———. 1979a. "Individualism and the Mental." *Midwest Studies in Philosophy* 4: 73–121.

———. 1979b. "Sinning Against Frege." *Philosophical Review* 88: 398–432.

———. 1982a. "Other Bodies." In *Thought and Object: Essays on Intentionality*, edited by Andrew Woodfield, 97–120. Cambridge: Oxford University Press.

———. 1982b. "Two Thought Experiments Reviewed." *Notre Dame Journal of Formal Logic* 23: 284–93.

———. 1986a. "Intellectual Norms and Foundations of Mind." *Journal of Philosophy* 83: 697–720.

———. 1986b. "Individualism and Psychology." *Philosophical Review* 95: 3–45.

———. 1989. "Wherein is Language Social?" In *Reflections on Chomsky*, edited by Alexander George, 175–91. Oxford: Basil Blackwell.

Canfield, John. 1983. "Discovering Essence." In *Knowledge and Mind: Philosophical Essays*, edited by Carl Ginet, 105–29. New York: Oxford University Press.

Crane, Tim. 1996. "All the Difference in the World." In Pessin and Goldberg, 284–304.

Donnellan, Keith. 1983. "Kripke and Putnam on Natural Kind Terms." In *Knowledge and Mind: Philosophical Essays*, edited by Carl Ginet, 84–104. New York: Oxford University Press.

Goosens, William. 1996. "Underlying Trait Terms." In Schwartz, 133–54.

Hanna, Robert. 1998. "A Kantian Critique of Scientific Essentialism." *Philosophy and Phenomenological Research* 58, no. 3: 497–528.

Kripke, Saul. 1972. *Naming and Necessity*. Cambridge: Harvard University Press.

Linsky, Bernard. 1998. "Putnam on the Meaning of Natural Kind Terms." *Canadian Journal of Philosophy* 74: 819–28.

Loar, Brian. 1988. "Social Content and Psychological Content." In *Contents of Thought*, edited by H. Grimm and D. D. Merrill, 99–110. Tucson: The University of Arizona Press.

McKay, T., and C. Stern. 1979. "Natural Kind Terms and Standards of Membership." *Linguistic and Philosophy* 3: 27–34.

McCulloch, Gregory. 1995. *The Mind and Its World*. London: Routledge.

McDowell, John. 1996. "Putnam on Mind and Meaning." In Pessin and Goldberg, 305–17.

———. 1994. *Mind and World*. Cambridge: Harvard University Press.

———. 1998. "The Woodbridge Lectures, 1997: 'Having the World in View: Sellars, Kant and Intentionality'." *The Journal of Philosophy* 95, no. 9: 431–91.

Mellor, D. H. 1996. "Natural Kinds." In Pessin and Goldberg, 69–80.

Pessin, A. and S. Goldberg, eds. 1996. *The Twin Earth Chronicles*. Armonk: N.Y.: M. E. Sharpe,

Poncinie, Lawrence. 1985. "Meaning Change for Natural Kind Terms." *Nous* 19: 415–27.

Putnam, Hilary. 1962. "It Ain't Necessarily So." *The Journal of Philosophy* 59: 658–70.

———— . 1970. "Is Semantics Possible?" In *Mind, Language and Reality, Philosophical Papers* II. Cambridge: Cambridge University Press, 139–52.

———— . 1973. Meaning and Reference." *Journal of Philosophy* 70: 699–711.

———— . 1975. "The Meaning of 'Meaning'." In *Mind, Language and Reality, Philosophical Papers* II Cambridge: Cambridge University Press, 215–71.

———— . 1978. "Reference and Understanding." In *Meaning and the Moral Sciences*. Boston: Routledge & Kegan Paul.

———— . 1981. "A Problem About Reference." In *Reason, Truth, and History*. Cambridge: Cambridge University Press.

———— . 1982. "Comments on Fodor's 'Cognitive Science and the Twin Earth Problem'." *Notre Dame Journal of Formal Logic* 23: 294–95.

———— . 1986. "Meaning Holism." In *The Philosophy of W. V. Quine*, edited by L. E. Hahn and P. A. Schilpp, 405–31. La Salle: Open Court.

———— . 1987. "Meaning Holism and Semantic Holism." In *Theorie der Subjektivitat*, edited by Konrad Cramer *et al*. Frankfurt am Main: Suhrkamp, 251–77.

———— . 1992. "Is It Necessary that Water is H_2O?" In *The Philosophy of A. J. Ayer*, edited by Lewis Edwin Hahn, 429–54. Peru: Open Court.

Stephen Schwartz, ed. 1977. *Naming, Necessity and Natural Kinds*. Ithaca: Cornell University Press.

———— . 1996. "Putnam on Artifacts." In Pessin and Goldberg, 81–88.

Kim Sterelny. 1996. "Natural-Kind Terms." In Pessin and Goldberg, 98–114.

Van Brakel, J. 1990. "Units of Measurement and Natural Kinds: Some Kripkean Considerations." *Erkenntnis* 33: 297–317.

Zemach, Eddy. 1996. "Putnam's Theory on the Reference of Substance Terms." In Pessin and Goldberg, 60–68.