

Naming Natural Kinds

Åsa Maria Wikforss
Stockholm University
Department of Philosophy
106 91 Stockholm
asa.wikforss@philosophy.su.se

Naming Natural Kinds

Can it be known a priori whether a particular term, such as ‘water’, is a natural kind term? Or is this something that cannot be known in advance of empirical investigations? If the latter is true, what are the implications for the doctrine, as famously presented in Putnam’s Twin Earth thought experiment, that natural kind terms require an externalist semantics (Putnam 1975a)?

The traditional assumption has been that we can distinguish natural kind terms from other types of terms a priori, on the basis of how we use these terms. According to this conception of natural kind terms, whether a term ‘T’ is a natural kind term depends solely on the intrinsic semantic features of ‘T’ (see Goosens 1977 and McKinsey 1987). This view accounts well for the intuition that natural kind terms, like proper names, are to be singled out for a special semantic treatment. However, it has difficulties accounting for the fact that it sometimes turns out that a purported natural kind term fails to name a natural kind. Familiar examples are ‘air’, ‘sand’, and ‘jade’. For this reason, an alternative conception of natural kind terms has been proposed and is currently gaining territory. Whether ‘T’ is a natural kind term, on this conception, depends on whether in fact it succeeds in picking out a natural kind, not on any a priori knowable semantic features of ‘T’. However, it is held, this does not have any implications for the externalist account of natural kind terms. It is possible to grant that it is an a posteriori question whether ‘T’ is a natural kind term, it is argued, and yet hold on to the view that natural kind terms are to be given an externalist semantics (see Boghossian 1997, Brewer 2000, Gallois 1996, McLaughlin and Tye 1998a and 1998b).

In this paper I shall argue that there are indeed strong reasons to reject the traditional conception of natural kind terms. Contrary to received opinion, however, I hold that this fact raises serious problems for externalism. Putnam's Twin Earth thought experiment, I argue, relies on the assumption that it can be known a priori that 'water' is a natural kind term. Thus, a choice must be made: Either hold on to the traditional conception of natural kind terms, with its inherent problems, or reject the idea that natural kind terms should be given an externalist semantics. Given this choice, I suggest we should take the latter option and reject externalism.

Through the years a number of objections have been raised to the Putnamian account of natural kind terms. In particular, it has been argued that the standard notion of a natural kind, relied on by the externalist, is problematic. What characterizes a natural kind on the standard view is that there is one underlying microstructural property, rather than a motley, uniting and explaining certain macrolevel properties (see for instance Putnam 1970, Sterelny 1983, and Wilkerson 1998). The objection is that the very idea of a common underlying microstructural property is spurious, and runs into difficulties of various sorts, such as the problems of isotopes and the problem of impurities (in Mellor 1977 and Zemach 1976).¹ However, I shall waive these difficulties and assume, for the sake of argument, that the externalist has a good reply to them.² That is, I will assume that the standard notion of a natural kind can be defended. Even given this assumption, I argue, the externalist account of natural kind terms is in trouble.

In my discussion I shall rely on the notion of a priori knowledge that is typically employed in the current debate, according to which a proposition is a priori knowable if it can be justified independently of empirical investigations (see Boghossian 1996, 161, and McLaughlin and Tye 1998a, 286). It is sometimes

suggested that a proposition can be a priori, in this sense, and yet be falsified by future empirical discoveries. I shall leave that issue aside however, since, I will argue, the externalist needs to make the stronger assumption, that future empirical discoveries cannot falsify the claim that a given term is a natural kind term.

The structure of the paper is as follows. Section 1 provides a brief account of Putnam's original discussion of the semantics of natural kind terms and the motivations behind his externalist treatment of these terms. Section 2 examines a version of the traditional account of natural kind terms, due to Michael McKinsey, and suggests that it is highly problematic. Section 3 argues that once it is granted that it is a purely a posteriori, empirical question whether a term names a natural kind, externalism is undermined. Section 4 considers three possible replies to my argument, but finds them all wanting. In particular, it considers the proposal that natural kind terms are to be given a 'disjunctive semantics'. The paper concludes by suggesting that the time has come to give a descriptivist account of natural kind terms a second hearing.

1. Putnam on Natural Kind Terms

It is not an accident that Putnam's first discussion of externalism focuses on natural kind terms. In several papers prior to "The Meaning of 'Meaning'" Putnam expresses skepticism about the prospects of giving a traditional semantic account of these terms. For instance, in "Is semantics possible?" Putnam singles out natural kind terms for special attention. After noting that there is a great variety of general terms, Putnam writes:

An important class, philosophically as well as linguistically, is the class of general names associated with natural kinds – that is, with

classes of things that we regard as of explanatory importance: classes whose normal distinguishing characteristics are ‘held together’ or even explained by deep-lying mechanisms. Gold, lemon, tiger, acid, are examples of such nouns (Putnam 1970, 139).

Putnam’s claim is that traditional theories of meaning radically falsify the properties of these nouns. According to the traditional view, Putnam says, the meaning of ‘lemon’ would be given by specifying a set of properties $P_1 \dots P_n$. For each of these properties, the statement ‘lemons have the property P’ expresses an analytic truth, and if something has all the properties listed it is necessarily a lemon. However, what characterizes natural kind terms, Putnam argues, is precisely that no such list of necessary and sufficient properties can be given. Whatever properties we list as defining characteristics of a lemon (yellow color, tart taste, a certain kind of peel, etc.) it is quite possible for something to be a lemon without having one of these properties (a green lemon is still a lemon) and from the fact that something has all these properties it does not follow that it is a lemon.

In order to account for this peculiar quality of natural kind terms, Putnam appeals to the idea that we use natural kind terms to pick out an underlying ‘essential nature’. To describe something as a *lemon*, Putnam says, is to indicate that it is likely to have certain characteristics, but also that these characteristics are to be accounted for by some ‘essential nature’ that this lemon shares with other members of the natural kind (Putnam 1970, 140). It is this possession of an underlying essential nature that allows for the possibility that something may have all the superficial properties associated with lemons and not be a lemon. It follows, Putnam suggests, that the traditional view that meaning determines extension must be rejected in the case of natural kind terms: “Meaning does not determine extension, in the sense that given the

meaning and a list of all the ‘properties’ of a thing...one can simply read off whether the thing is a lemon (or acid, or whatever)” (Putnam 1970, 142).

It is clear that there is a short step from these observations to Putnam’s claim in “The Meaning of ‘Meaning’” that “meanings ain’t in the head”. If no traditional definition can be given of a natural kind term, if no set of necessary and sufficient properties can be listed, then the meaning of a natural kind term cannot be exhausted by the set of properties a speaker associates with it. To illustrate his point Putnam designs a thought experiment that is, by now, so familiar that a very brief description will suffice. Putnam imagines two planets, Earth and Twin-Earth that are exactly the same, except that the liquid called ‘water’ on Twin Earth has a very different chemical composition, XYZ for short, than the liquid called ‘water’ on Earth, which has the chemical composition H₂O. Despite this difference in microstructural make-up, water and twin-water share all macro-level properties, and before the advance of modern chemistry it would have been impossible to tell them apart. The speakers on Earth therefore associate all the same macro-level properties with their term ‘water’ as the speakers on Twin-Earth associate with their term ‘water’.³ Nonetheless, Putnam argues, the difference in underlying chemical composition implies that ‘water’ has a different extension on the two planets and, as a consequence, a different meaning: “If a spaceship from Earth ever visits Twin Earth, then the supposition at first will be that ‘water’ has the same meaning on Earth and on Twin Earth. This supposition will be corrected when it is discovered that ‘water’ on Twin Earth is XYZ...” (Putnam 1975a, 223)

Just as in the earlier paper, Putnam puts his point by saying that meaning fails to determine extension in the case of natural kind terms. This, again, is intended as a refutation of the traditional descriptivist account of reference-determination.

However, as Putnam himself makes clear, in another sense he does accept the claim that meaning determines extension: i.e. he holds on to the principle that a difference in extension makes for a difference in meaning. This principle is essential to Putnam since it allows him to move from the claim that there is a difference in the extension of 'water' on the two planets to the conclusion that there is a difference in meaning.⁴ As Putnam himself points out, the principle does not hold for pure indexicals. For instance, 'I' shifts reference from one context of utterance to another but it does not therefore shift meaning. However, Putnam argues, natural kind terms cannot plausibly be compared to pure indexicals (Putnam 1975a, 245-246). Consequently, according to Putnam, if 'water' has a different extension on Twin earth than on Earth, it follows that it has a different meaning on the two planets.

Given this stage setting, it is time to turn to the question of how Putnam's externalist account of natural kind terms relates to the two conflicting conceptions of natural kind terms mentioned in the introductory remarks.

2. Natural Kind Terms and the A Priori: A Semantic Approach

One of the most extensive defenses of the traditional idea that it is a priori knowable whether a term is a natural kind term can be found in the writings of Michael McKinsey (McKinsey 1987 and 1994). It is instructive to have a closer look at McKinsey's account, since the explicit motivation behind McKinsey's account is to give a semantics of natural kind terms that supports Putnam's externalism.

McKinsey raises the question why natural kind terms should be given an externalist construal or, as he puts it, why these terms express 'wide' properties. It is quite clear, he argues, that proper names and indexicals must be understood externalistically, but it is less clear why natural kind terms should be understood that

way. Natural kind terms clearly have linguistic meanings in the public language, unlike proper names -- it is not their sole function to pick out an object. And unlike indexicals natural kind terms do not change reference from one utterance context to another. Why then, should it be said that natural kind terms express wide properties? For instance, why should we draw the conclusion that 'water' on Twin Earth has a different meaning than 'water' on Earth?

The answer, McKinsey suggests, is that natural kind terms express *relational* predicates, i.e. they are used to predicate membership in a particular natural kind. The predicate 'is water', McKinsey says, is used to ascribe the "purely relational property of belonging to a certain natural kind, and the linguistic meaning of 'water' is given by a semantic rule that determines the property in question" (McKinsey 1994, 322). The rule, in turn, is an 'objectual rule', a rule which contains an essential reference to *us*, the speakers using the term. Thus, 'water' has a different meaning on Twin Earth than on Earth, since 'water' expresses the relational property of belonging to that natural kind which we pick out by 'water' on our planet, i.e. H₂O and not XYZ.

According to McKinsey, therefore, whether a particular term is a natural kind term depends solely on the semantic rule that governs the use of the term in question, not on whether it ultimately *succeeds* in picking out a natural kind. Moreover, since McKinsey holds that semantic rules are internalised by speakers, he also holds that the speaker can know a priori whether a particular term is a natural kind term (McKinsey 1987, 5). The advantage of McKinsey's view to the externalist is that it provides a semantics of natural kind terms that gives straightforward support to Putnam's claim that 'water' has a different meaning on Twin Earth than on Earth. It is because it is part of the meaning of 'water' that its semantic function is to predicate membership in a natural kind that the difference in the chemical composition of the liquids on Earth

and on Twin Earth has semantic significance. However, there is a clear drawback to a view of McKinsey's sort.

The question is what should be said in those cases where it is discovered that a purported natural kind term fails to actually pick out a natural kind. As mentioned above, history is full of examples of terms that once were thought to pick out a natural kind but were later discovered not to do so.⁵ That these discoveries are made is hardly surprising. What characterizes a natural kind, on the standard view, is that there is one underlying microstructural property, rather than a motley, uniting and explaining certain macrolevel properties, and it clearly cannot be known a priori that there is one such underlying property. Indeed, it seems easy to imagine scenarios in which we would conclude that 'water' does not succeed in picking out a natural kind. Putnam himself allows for this. If we were to discover that XYZ, in addition to H₂O, is plentiful on Earth, Putnam says, "then we would have had a case similar to the jadeite/nephrite case: it would have been correct to say that there were *two kinds of 'water'*" (Putnam 1975a, 241). That is, if we were to discover that there is a liquid superficially just like water but with a different underlying chemical composition, one that equally well explains the superficial qualities we use to identify water, the conclusion to be drawn is not that this liquid is not water, but that 'water' is not a natural kind term. The conclusion would be that our belief that water is a natural kind is false, just as our belief that jade is a natural kind turned out to be false.

On McKinsey's view, however, we cannot say this, since, for him, 'is water' expresses the relational property of belonging to a certain natural kind. In the scenario described above the rule for 'water' would fail to pick out a property of the required sort (since there is no such natural kind) and so, according to McKinsey, any statement about water would fail to express a proposition, it would lack truth-value.

Thus, if ‘water’ turns out not to name a natural kind it would not be true that water is wet, that there is water in the oceans, that water quenches thirst, etc., and our discovery could not be described as the realization that our belief that water is a natural kind is false.⁶

The difficulty with this approach is not limited to McKinsey’s particular account, but applies to any account according to which it is part of the meaning of certain terms that they predicate membership in a natural kind.⁷ Putnam, in his 1970 discussion of natural kind terms anticipates this difficulty. After having argued that we cannot give an ordinary definition of natural kind terms, since a natural kind may have abnormal members, Putnam tries out the following definition: “X is a *lemon* = *df.* X belongs to a natural kind whose normal members have yellow peel, tart taste, etc.” Putnam then goes on to argue that although this definition would allow for abnormal lemons, it has a fatal flaw. The definition implies that if there is no natural kind whose normal members are yellow, tart, etc., “ then even these tart, yellow, thick-peeled fruits that I make lemonade from are *not literally lemons*” (Putnam 1970, 142). But, Putnam says, this is absurd and so the suggested definition must be rejected.

Although, therefore, a view of McKinsey’s sort supports the externalist conclusions of Putnam’s thought experiment, it has the counterintuitive implication that discovering that a purported natural kind term fails to name a natural kind entails a reference failure. The interesting question, then, is whether we can give up the assumption that it is part of the meaning of certain terms that they predicate membership in natural kinds without thereby undermining the externalist account of natural kind terms. This is the question that will be pursued in the remainder of the paper.

3. *Fallible Intentions and Twin Earth*

It appears that there is a simple reply to this question. The externalist, it might be argued, need not make the implausible claim that it is part of the *meaning* of ‘water’ that it names a natural kind. It is sufficient to appeal to the speaker’s *intention* that ‘water’ names a natural kind. The externalist can then say that there is a sense in which it can be known a priori that ‘T’ is a natural kind term, i.e. it can be known a priori that ‘T’ is intended to pick out a natural kind. This, it can be argued, suffices to secure the conclusions of Putnam’s thought experiment, while allowing for the possibility that a purported natural kind term does not succeed in picking out a natural kind. After all, intentions can fail.

It is plausible that Putnam had something like this in mind in “The Meaning of ‘Meaning’”. Thus, to support his externalist conclusions he explicitly appeals to our intentions when using terms like ‘water’. The reason we should conclude that ‘water’ has a different meaning on Twin Earth than on Earth, he argues, is that when we ostensively defined ‘water’ on Earth, we *intended* ‘water’ to be used as a natural kind term -- we intended it to pick out that substance which underlies our water-samples (Putnam 1975a, 231 and 238). At the same time Putnam holds that this intention can fail since, as noted above, he wants to allow for the possibility that water is not a natural kind, but more like jade.

Plausible as this suggestion may seem, however, it does not work. The difficulty is this: How can the speaker’s intention that ‘water’ names a natural kind bear the semantic weight required if it is also allowed that this intention can fail? If the intention can fail, would not the discovery of XYZ on Twin Earth be precisely the kind of discovery that would show one’s intention to have failed? The discovery of XYZ, after all, would show that there is more than one chemical composition that

underlies and explains all of the superficial qualities that we associate with ‘water’. The conclusion to be drawn when encountering Twin Earth, then, would not be that ‘water’ has a different meaning than on Earth, but that water is not a natural kind, that there are two kinds of water (see Häggqvist 1996, 171, for a related suggestion).

To bring out the problem, recall Putnam’s original motivation for saying that natural kind terms defy traditional semantics. Whatever observable properties we list as defining the term ‘lemon’, Putnam claims, something may have all these properties and yet fail to be a lemon, since ‘lemon’ is used to pick out an essential underlying property. However, if it is also held that it is a *falsifiable* assumption that ‘lemon’ names a natural kind this claim is undermined. That is, if the intention that ‘lemon’ names a natural kind can fail, then the discovery of a fruit with all the observable properties of ordinary lemons, but with a different microstructural make-up, would constitute a discovery that lemons do not form a natural kind.

The difficulty also crops up in Kripke’s discussion of natural kinds. Consider his discussion of species, such as tigers. Even though we don’t *know* the internal structure of tigers, Kripke suggests, we suppose that tigers form a certain species or natural kind. Thus, he continues, “[w]e can say in advance that we use the term ‘tiger’ to designate a species, and that anything not of this species, even though it looks like a tiger, is not in fact a tiger” (Kripke 1972, 121). At the same time, however, Kripke grants that our supposition that tigers form a natural kind may be mistaken. But if this is so, quite clearly, we *cannot* say in advance that an animal that looks like a tiger but has a different internal structure is not a tiger. On the contrary, if we encounter animals indistinguishable from normal tigers but with a different internal structure, we should conclude not that the animal is not a tiger but that our supposition that tigers

form a natural kind is falsified. Our original supposition therefore cannot carry the kind of semantic significance that Kripke wishes to ascribe to it.⁸

The trouble, in short, is that the speaker's intention that the term in question name a natural kind cannot be appealed to in order to support Twin Earth-type thought experiments if it is also allowed that the intention can fail. Indeed, it is odd to speak of an *intention* here at all, since it is odd to say that one can intend something that one has no control over.⁹ Rather, what should be said is that the speaker has the *belief* that 'water' names a natural kind, a belief that may turn out to be false.

It therefore appears that the externalist account of natural kind terms presupposes a view of McKinsey's sort. If this is so, the externalist is saddled with the implausible claim that if 'water' fails to name a natural kind we have a case of reference failure. But perhaps this is just an appearance? Next I will consider three replies commonly given on behalf of the externalist – replies that all aim to avoid a view of McKinsey's sort and yet preserve the externalism.

4. Three Replies

4.1 The importance of local water

The first reply appeals to the idea that there is all the difference between discovering XYZ on Twin Earth and discovering XYZ on Earth. The discovery of XYZ on Earth, it is argued, would indeed constitute a falsification of the assumption that 'water' names a natural kind, whereas the discovery of XYZ on Twin Earth would not. This is so since the speaker's intention is to pick out the substance that underlies the *local* 'water'-samples, i.e. the samples on Earth. In fact, Putnam says this much by arguing that natural kind terms have an unnoticed indexical component. 'Water', he says, "is

stuff that bears a certain similarity relation to the water *around here*” (Putnam 1975a, 234).

But this reply is not convincing. If ‘water’ can fail as a natural kind term, why should only samples taken on planet Earth matter? What if we find XYZ on Mars? Or is it just our solar system that matters? But any of these parochialisms seem entirely arbitrary. It is often stressed that the category of natural kind terms is essential to the inductive practices of the natural sciences, but no such practice limits itself to planet Earth. It is of course a different matter if, like McKinsey, one builds into the meaning of ‘water’ that it picks out a natural kind on the planet that the speakers happen to inhabit, but those who deny that it is part of the meaning of ‘water’ that it names a natural kind cannot similarly motivate the focus on a particular planet.

It might be said that this misses the point of referring to ‘local’ samples. The point, it might be objected, is not that *planet Earth* is particularly important, but that the *causal history* of our use of the term is: It is because our use of ‘water’ is causally anchored in H₂O, rather than XYZ, that XYZ falls outside the extension of ‘water’. If, instead, we were frequently exposed to both XYZ and H₂O (if, say, we traveled back and forth between Earth and Twin Earth) then ‘water’ would be causally grounded in both of these chemical compositions, and it would not be a natural kind term. Thus, by appealing to the causal history of the term’s use, it is possible both to secure the conclusion that XYZ is not water, and allow for the possibility that ‘water’ fails to name a natural kind, without making an arbitrary appeal to planet Earth.

However, this reply does not avoid the problem of arbitrariness. Consider the case in which there is plenty of both XYZ and H₂O around, but I happen to have causally interacted only with H₂O (see Fodor 1994, 32, for a discussion of this case). Does this mean that my word is ‘causally anchored’ in H₂O rather than XYZ? If so,

it would follow that no future encounter with XYZ could falsify the hypothesis that ‘water’ names a natural kind. But if it is a truly empirical hypothesis that ‘water’ does name a natural kind, then a future falsification of this sort must be allowed for; indeed, it would seem that that is exactly what it means for it to be an empirical hypothesis that ‘water’ names a natural kind. To allow for future falsifications, without it following that Twin Earth constitutes such a falsification, the defender of the ‘causal anchoring’-picture has to argue that only *certain kinds* of future falsifications are permitted such that finding XYZ in one’s immediate environment would constitute a falsification, but finding XYZ somewhere else would not. But this just takes us back where we were: If ‘water’ can fail as a natural kind term, why should it matter *where* XYZ is found?

4.2 The Counterfactual Scenario

This invites another reply on the part of the externalist. In the discussion above it has been assumed all along that Earth and Twin Earth are in the *same* world. It is because of this assumption that it could be argued that the discovery of twin water would constitute a falsification of the hypothesis that ‘water’ is a natural kind term.

However, it might be objected, although Putnam did place the two planets in the same world, the best route to externalism is by placing the planets in different possible worlds. The externalist conclusions can then be reached without relying on the assumption that it is part of the meaning of ‘water’ that it predicates membership in a natural kind. Thus, it could be argued, if in fact ‘water’ picks out one underlying chemical composition in the actual world then it is a natural kind term, and any world in which ‘water’ does not pick out that composition is one in which ‘water’ must have

a different meaning. This, it might be said, is just the strategy employed by Kripke to produce a posteriori necessities.

It should be noted, at the outset, that this move in itself constitutes a major concession on part of the externalist. The general assumption has been that Putnam's Twin Earth-experiment can be set in the actual world, without having to rely on intuitions about counterfactual worlds. Indeed, it is clear that although Putnam does endorse some of Kripke's vocabulary in "The Meaning of 'Meaning'", he deliberately avoids an appeal to counterfactual worlds when setting up the thought experiment. In a later paper Putnam himself emphasizes this point. He stresses that in "The Meaning of 'Meaning'" possible worlds were mentioned only in connection with Kripke, and that Twin Earth were to be understood as a far-away planet in the actual world:

We were to imagine that a liquid on this Twin Earth superficially resembles water. How do we decide whether it really *is* water? Since the question only concerned *actual* substances, questions about 'all possible worlds' – in particular questions about worlds in which *the laws of nature can be different* – were not in my mind" (Putnam 1990, 69).

Putnam then goes on to reject the very idea of a criterion of substance-identity that extends to possible worlds, and argues that the question 'What is the necessary and sufficient condition for being water *in all possible worlds?*' makes little sense (ibid. 70).

It is therefore not a trivial matter if Putnam's thought experiment has to be set across possible worlds. However, even if we ignore these scruples about cross-world substance identity,¹⁰ it is an illusion to think that placing the two planets in different worlds saves the externalist from having to make the assumption that it is part of the

meaning of 'water' that it predicates membership in a natural kind. There are two, related, difficulties.

First, it was noticed above that a central premiss for Putnam's argument that 'water' has a different meaning on Twin Earth, is the principle that a difference in extension makes for a difference in meaning (in the case of non-indexicals). However, quite clearly, this principle does not apply *across worlds*. After all, it is a trivial truth that there are possible worlds in which our word 'water' has a different extension than in the actual world, and yet has the same meaning as in the actual world (just imagine a counterfactual world exactly like ours minus Lake Michigan).¹¹ Thus, once the planets are set in different possible worlds, it cannot be concluded that 'water' has a different meaning on Twin Earth than on Earth simply because 'water' has a different extension on the two planets. Rather, what has to be shown is that the difference in extension is such that it warrants the conclusion that there is a difference in meaning, and it is unclear how this can be shown unless it is assumed to be part of the meaning of 'water' that it predicates membership in a natural kind.

The second, and related, difficulty concerns the appeal to Kripke's strategy for generating a posteriori necessities. Metaphysical necessities, as is often pointed out, cannot be generated in a vacuum but depend on semantic presuppositions (see Forbes 1997, 529; Hale 1997, 492; Salmon 1982, 184; Sidelle 1989). It may be that it is possible to conclude that water is necessarily H₂O (if it is H₂O), without assuming that it is part of the meaning of 'water' that it is H₂O, but it is not possible to reach such a conclusion without making *any* assumptions about the meaning of 'water'. After all, the fact that 'water' and 'H₂O' happen to be co-extensional in the actual world cannot force the conclusion that they are co-extensional in all possible worlds. What has to be added, as Kripke emphasizes, is that 'water' is a rigid designator, and

the motivation for this is precisely the assumption that ‘water’ is a natural kind term. In short, Kripke’s strategy presupposes that it is intrinsic to the semantics of ‘water’ that it is a natural kind term.¹²

It might be objected that this expresses a misunderstanding, since the very point of Kripke’s suggestion that ‘water’ is a rigid designator, is to reject an appeal to descriptive meanings. Descriptions, on Kripke’s view, serve to fix reference, but that does not make them ‘part of the meaning’ (Kripke 1972, 57). Thus, we can use certain descriptions to fix the reference of ‘water’ (thirst-quenching, transparent, etc.) without it following that any of them are unrevisable or intrinsic to the semantics of ‘water’.

This, of course, is the picture Kripke sketches, but the question is whether the assumption that ‘water’ names a natural kind can be treated the same way as these other descriptions. After all, the argument for why these other descriptions are not part of the meaning of ‘T’ relies on the assumption that ‘T’ is a natural kind term. For instance, when arguing that something might have all the identifying marks of gold and yet not be gold, Kripke appeals to the idea that we use ‘gold’ as a term for a certain kind of thing, a certain substance (Kripke 1972, 118-119). If the assumption that ‘gold’ is a natural kind term could be revised, then Kripke could not argue on purely a priori grounds, the way he does now, that ‘gold’ is a rigid designator and cannot be given a descriptive account.¹³

4.3 A Disjunctive Solution?

It might be thought that an important option has been over-looked in the above discussion. This is simply to give a *disjunctive* construal of the semantics of natural kind terms. On this view, ‘water’ names that microstructural property that underlies and explains certain superficial properties, *if there is such a property*, otherwise it

names the superficial kind. This, it could be argued, is sufficient to ensure the result that if indeed ‘water’ names a natural kind then this goes in to the individuation of the meaning of ‘water’, and yet the undesirable consequences of McKinsey’s position are avoided since the possibility that ‘water’ names a superficial kind is allowed for.

This reply appears to have a great deal of appeal among current-day externalists.¹⁴ It is not hard to understand its intuitive appeal. After all, if we grant that it is an empirical assumption that ‘water’ names a natural kind, as I have argued that we should, then it follows that either ‘water’ names a natural kind or it names a superficial kind. This is just a truism. However, it is clear, the externalist cannot settle with this truism, since the suggestion that either ‘water’ names a natural kind or a superficial kind in itself fails to imply any externalist thesis. Indeed, as we shall see, the suggestion is perfectly compatible with a descriptivist account. What the externalist has to add, is that we use ‘water’ in such a way that if indeed it names a natural kind, *then this affects the semantics of the term*, ruling out a descriptivist account.

To illustrate this view, let us consider a recent account of Twin Earth-externalism, defended by Brian McLaughlin and Michael Tye (1998a, 297-299; 1998b, 370-372). McLaughlin and Tye’s main concern is with the question whether externalism is compatible with self-knowledge, and they set out to respond to a well-known reductio argument, first put forth by McKinsey, against the compatibilist position (McKinsey 1991). McKinsey’s argument, in brief, is that if one assumes content externalism to be compatible with the idea that we have a priori knowledge of the content of our own thoughts, it follows that one could have a priori knowledge of facts about the external world that no one would say are knowable that way. For instance, if my thoughts about water are externalistically individuated, and I could

know a priori that I am thinking the thought *water is a liquid*, it follows, absurdly, that I could know a priori that there is water in my environment (see also Boghossian 1997).

In response to this, McLaughlin and Tye reject the assumption that the externalist is committed to the idea that if it can be known a priori that one is thinking that water is a liquid, it can be known a priori that there is water in the environment. What the externalist is committed to, they suggest, is merely the hypothetical that *if* ‘water’ names a natural kind, *if* ‘water’ expresses a natural kind concept, then this serves to individuate the individual’s ‘water’-thoughts. However, it cannot be known a priori that ‘water’ names a natural kind since, they argue, “water might have turned out to be like air, or like jade” (McLaughlin and Tye 1998b, 370). In such a case, they suggest, the concept expressed by ‘water’ would have to be understood descriptively, in terms of its conceptual role. Consequently, the fact that the individual can know a priori that he has the concept of water does not imply that he knows a priori that there is water in the environment. A crucial piece of a posteriori knowledge must also be added, i.e. that the concept of water is a natural kind concept. McLaughlin and Tye write: “while one can know a priori that one is thinking that water is a liquid, one cannot know a priori that one is having a thought involving a natural kind concept; for one cannot know a priori that the concept of water succeeds in picking out a natural kind” (ibid. 371).

McLaughlin and Tye’s claim is therefore not simply that it is an a posteriori question whether ‘water’ is a natural kind term. Rather, the claim is that it is an a posteriori question what *semantics* ‘water’ is to be given: Either ‘water’ names a natural kind, in which case it is to be given an externalist semantics, or it names a superficial kind, in which case it can be given a traditional descriptivist semantics. Or,

to put it in terms of concepts: Either the concept of water succeeds in picking out a natural kind, in which case it is externalistically individuated, or it does not, in which case it is individuated descriptively, in terms of its conceptual role (ibid. 371; see also McLaughlin and Tye 1998a, 305-307). André Gallois, who has defended a similar position, calls it ‘the semantic variation thesis’. This is very apt, since, on this view, what varies with the external environment is indeed the semantics of the term given. Discussing the case of ‘phlogiston’, Gallois suggests that if phlogiston exists as a natural kind then the term ‘phlogiston’ is a rigid designator, whereas if phlogiston does not exist then ‘phlogiston’ functions differently: “For example, if phlogiston does not exist, ‘phlogiston’ is a synonym for a non-rigid definite description such as ‘the stuff thought by some eighteenth-century chemists to have negative weight’” (Gallois 1996, 186).

It should be clear how radical a position this is, and how far removed it is from Putnam’s and Kripke’s original discussions. Putnam and Kripke both reach their anti-descriptivist conclusions by arguing that intuition tells us that none of the descriptions associated with ‘water’ or ‘gold’ are necessary. On the disjunctivist view, however, intuition can tell us nothing about the modal status of statements about water, since intuition can tell us nothing about the semantics of ‘water’. If it cannot be known a priori whether a term ‘T’ is to be given a descriptive semantics or not, then it cannot be known a priori whether the descriptions associated with ‘T’ merely serve to fix the reference of ‘T’, or whether they are part of the meaning of ‘T’. Thus, if ‘water’ names a natural kind, then the descriptions associated with ‘water’ merely serve a reference-fixing function and are all contingent; if ‘water’ does not name a natural kind, by contrast, the descriptions are part of the meaning of ‘water’ and cannot be revised. As a result, all modality becomes a posteriori – metaphysical as well as

conceptual. Kripke's and Putnam's reasoning is therefore fundamentally flawed: To determine what semantics 'water' is to be given we cannot rely on a priori reasoning but need to undertake empirical investigations.

By the same token, no thought-experiment, such as the Twin Earth-experiment, can be used to support the a posteriori externalist position. McLaughlin and Tye write that the Twin Earth thought experiment relies on the *empirical* assumption that water is a natural kind and that "such thought experiments are *not* intended to show that the concept of water is a natural kind concept" (McLaughlin and Tye 1998b, 370). It is of course correct, as argued above, that Putnam's thought experiment is not intended to *show* that 'water' is a natural kind term, but presupposes that it is. What the thought experiment *is* supposed to tell us, however, is something about the meaning of *our* term 'water', something about *our* concept of water; namely, that it cannot be given a descriptivist account. However, if it is an *empirical* presupposition that 'water' is a natural kind term, then the experiment provides no evidence one way or the other. At the most, the experiment tells us something about a hypothetical language, a hypothetical concept (the natural kind concept *water*), but it cannot tell us whether this hypothetical language, this concept, in fact is ours -- determining that would require empirical investigations.

Disjunctivism therefore undermines the possibility of relying on a priori arguments of the sort Kripke and Putnam do in order to support their anti-descriptivist position. This makes the externalism curiously unmotivated. It also makes it difficult to understand how disjunctivism can seem appealing to somebody, like McLaughlin and Tye, who wishes to defend the view that externalism is compatible with the idea that we know the content of our own thoughts *a priori*, independently of empirical investigations. Having such a priori knowledge, one should think, implies having

some a priori access to the modal status of the thought, to its conceptual connections.¹⁵ After all, as Tyler Burge has argued, one reason we care about self-knowledge is its connection with critical reasoning, the ability to appreciate the rational connections between one's thoughts, and this ability is seriously impaired if I do not know whether 'water' expresses a descriptive concept or is a rigid designator (Burge 1996).

The trouble with disjunctivism is not merely epistemic, however. Even if the externalist is willing to bite the bullet and reject the reliance on a priori reasoning that characterizes Putnam's and Kripke's discussions, it remains obscure how the semantics of a term could be dependent on the external world in the way suggested by disjunctivism. Again, the disjunctivist is not merely saying, with Putnam, that the meaning of certain terms are individuated by the external environment. Rather, the claim is that the external environment determines what *semantics* a particular term is to be given. This transforms the externalist thesis from a thesis about the special linguistic function of some of our terms (such as natural kind terms and proper names) into a metaphysical thesis (potentially applicable to *any* term).¹⁶ *All* the semantic work is done by the external world, and the question is how the external world could play such a central semantic role. For instance, how could the chemical make-up of the world determine that 'water' is a rigid designator? Or the presence of nephrite that 'jade' is a non-rigid designator to be given a descriptivist account? In "The Meaning of 'Meaning'", Putnam emphasizes that the rigidity of 'water' follows from our linguistic intentions, and this is also how Kripke talks about rigid designators. For instance, Kripke emphasizes that the kind of rigidity he is concerned with is that of 'de jure' rigidity, where a designator is *stipulated* to be rigid.¹⁷ On the disjunctivist

view, however, rigidity is not a linguistic device, but flows from the world itself, and the disjunctivist owes us an account of how this could be so.

It is important to be clear that what I am objecting to is the idea of a disjunctive *semantics*. One reason it has seemed to many that a disjunctive account is unproblematic, no doubt, is the fact that we often speak of disjunctive *properties*. However, it is one thing to say that ‘water’ expresses a disjunctive property, quite another to say that it should be given a disjunctive semantics. To say that ‘water’ expresses a disjunctive property is just to say, for instance, that for something to be water it has to be H₂O *or* have a certain set of superficial qualities. This, of course, is precisely what the externalist does *not* wish to say, since if water were a disjunctive property of this sort XYZ would be water. The disjunctive externalist, as we have seen, is saying something very different; namely that if ‘water’ turns out to name a natural kind, then ‘water’ should be given an externalist semantics, otherwise ‘water’ can be understood descriptively. And it is this appeal to a disjunctive semantics that causes trouble, not the appeal to disjunctive properties.

5. Concluding Remarks

I have argued that the externalist account of natural kind terms depends on a highly problematic conception of natural kind terms, according to which a term is a natural kind term only if it is part of the term’s a priori knowable, semantic features, that it names a natural kind. If this is correct, we are left with a choice: The first option is to reject the claim that natural kind terms should be given an externalist semantics, and reconsider the possibility of giving a descriptivist account of these terms. The second option is to dig in our heels and hold on to the traditional conception of natural kind terms. Given this choice, I believe, the first option is much to be preferred. Although I

cannot here develop a descriptivist account of natural kind terms, let me end by indicating how the discussion above shows that a descriptivist account of natural kind terms is a lot more tenable than is commonly assumed.¹⁸

The clear advantage of descriptivism is that it allows us to give an unproblematic account of the intuition that if ‘water’ fails to name a natural kind, then it names a superficial kind instead. Consider, for instance, the traditional cluster theory, criticized by Putnam and others. On this view, the meaning of ‘water’ is determined by a cluster of descriptions, a cluster of beliefs associated with the term. Among this cluster are beliefs about the superficial properties of water, but also the belief that water is a natural kind; i.e. the belief that there is one underlying property uniting, and explaining the superficial set of properties associated with ‘water’.¹⁹ Although not all of these beliefs can turn out to be false, none of them express unrevisable, conceptual truths. Thus, the belief that water is a natural kind, although central to our conception of water, could turn out to be false -- for instance, if we were to discover twin-water. In such a case, the cluster theorist will simply say that ‘water’ is a superficial kind term. If there is no twin-water, however, then our belief that water is a natural kind stands strong, and ‘water’ is a natural kind term.²⁰

The descriptivist can say this, notice, without getting into the kinds of difficulties that plagues disjunctivism. On the descriptivist view, the fact that it cannot be known a priori whether ‘water’ is a natural kind term or a superficial kind term does not have any radical implications for our ability to reason a priori, since it is not held that if ‘water’ is a natural kind term, this affects the semantics of ‘water’. Rather, the semantics of ‘water’ is the same whether ‘water’ names a natural kind or not -- either way ‘water’ is to be given a descriptive account. It follows that the category of

natural kind terms, although interesting and important to science, does not form a special *semantic* category.²¹

Moreover, notice that once it is granted that it is an a posteriori question whether ‘T’ names a natural kind, the standard objection to descriptivism dissolves. The objection is found in Putnam’s discussion in “Is Semantics Possible?” and drives his thought experiment. According to Putnam, again, what characterizes natural kind terms is that whatever list of defining properties we give, something may have all the properties listed and yet not belong to kind K, as exemplified by twin-lemons and twin-water. The descriptivist cannot allow for this, Putnam continues, but is committed to saying that if something has all the listed properties then it belongs to kind K. For instance, if a liquid has all the observable features of water then, according to the descriptivist, it is water, even if it has a different microstructure, and this is wildly implausible, Putnam argues, given that ‘water’ is a natural kind term. We can now see where this objection goes wrong. It is of course quite right that if ‘water’ is a natural kind term then it cannot pick out several different chemical compositions in the same world. Indeed, this seems close to a definitional truth, at least on the standard conception of natural kinds. However, as should be clear by now, the descriptivist need not question this truth. Instead, the descriptivist should simply point out that it is a hypothetical claim that ‘water’ is a natural kind term, and that what Twin Earth shows is precisely that ‘water’ is *not* a natural kind term. Consequently, the possibility of twin-lemons, twin-water, etc. does not show that the descriptivist is committed to the implausible view that a natural kind term may pick out more than one underlying microstructure.

This also allows us to give an alternative explanation of the intuitive force behind Putnam’s Twin Earth externalism. Given the assumption that ‘water’ is a

natural kind term, it appears indisputable that ‘water’ has a different extension and a different meaning on Twin Earth. However, it should be clear by now, this reasoning involves a highly problematic simplification, as is often the case in thought experiments, since it simply ignores the question whether it can be known a priori that ‘water’ is a natural kind term. Once this question is raised, and it is granted that it cannot be known a priori whether ‘water’ is a natural kind term, the thought experiment comes in an entirely different light. It becomes possible to deny that ‘water’ has a different extension and meaning on Twin Earth, without, therefore, denying the truism that if ‘water’ is a natural kind term it names one underlying property only.

There are therefore reasons to believe that a descriptivist account of natural kind terms will fare a lot better than has standardly been assumed. Whether, in the end, descriptivism can be defended remains to be seen. What can be concluded is that the externalist doctrine is in sufficient trouble for a descriptivist account of natural kind terms to be given a second hearing.

¹ The problem of isotopes is discussed by Mellor 1977. Does the discovery that there are two common types of isotopes of chlorine show that chlorine is not a natural kind? If so, is there anything that appropriately could be called a ‘natural kind’? The problem of impurities is discussed in Zemach 1976. It derives from the fact that much of what we call ‘water’ is a mix of varying compositions – hard water, soft water, mineral water, salt water, etc. For a more recent criticism of the standard conception of a natural kind see Needham 2000. As Needham points out, chemical composition is distinct from microstructure, and although the chemical composition of water has long been known, little is in fact known about the microstructure of water. This is an

important point, but since it does not affect my argument I shall speak indifferently of chemical composition and microstructure

² For some recent replies see Abbott 1999, and Brown 1998.

³ Putnam sometimes puts this by saying that the speaker on Earth, Oscar, has exactly the same *concept* of water as his twin on Twin-Earth, Toscar. But this formulation is unfortunate, as subsequent discussion has shown. After all, if one takes concepts to be tied to reference, one will *not* grant that Toscar and Oscar have the same ‘water’-concept. Rather, one will take the thought experiment to show that concepts are individuated externalistically (see for instance Burge 1982 for this response). I therefore prefer to speak in terms of sameness of definitions and observable properties, leaving it open, at this point, whether the externalism extends to concepts and thought content.

⁴ This principle also plays an essential role in Tyler Burge’s defence of externalism, as Burge himself makes quite clear. See for instance Burge 1982, 107, and 1989, 186.

⁵ See Wilkerson 1993 for a discussion of the ubiquity of this phenomenon. Wilkerson emphasises that “many of our attempts to invent new natural-kind names misfire, because our assumption that there is some underlying property common to all the members of a supposed kind may be false.” (Wilkerson 1993, 5).

⁶ Scott Soames, in his recent book, discusses a view of McKinsey’s sort and suggests that it is implausible. We may presume, he says, that during the period before the discovery “speakers used sentences containing the term to convey lots of information” (Soames 2002, 282).

⁷ For related accounts see Brown 1998, 277, Goosens 1977, 49, and Jackson 1998, 46-52. Brown explicitly endorses the consequence that if a purported natural kind term fails to name a natural kind we have a case of reference failure. She says, for example,

that what ‘ruby’ refers to depends on whether ruby counts as a natural kind which in turn depends on whether mixtures are natural kinds: “My own view is that they are not, in which case ‘ruby’ would fail to have a reference.” (Brown 1999, 53)

⁸ In fact, there are reasons to suspect that tigers and other species are not natural kinds, on the standard view of natural kinds, since the genetic variations from one animal to another within a species, as well as the genetic continuity from one species to another, means that there is no good candidate for a ‘shared underlying structure’ of a particular species. For a discussion of this see Wilkerson 1993 and Wilson 1982, 553. It could be argued that this shows that there is something wrong with the standard notion of natural kinds (see for instance Dupré 1981). This may well be right but is not of any help to the externalist who relies on the standard notion.

⁹ What is odd is not intending things that in fact are impossible to obtain, but intending something one *knows* one has no control over. See Davidson 1992, 259, for this point.

¹⁰ Kripke certainly would do so, since he argues explicitly that we don’t need anything like a ‘criterion of transworld identity’ (Kripke 1972, 17). However, as Putnam emphasises, even if one accepts this in the case of names it is much less plausible in the case of general terms such as ‘water’ (Putnam 1990, 67).

¹¹ See Schwarz 1977, 38, and Häggqvist 1996, 177, for a discussion of this. Discussing Burge’s well-known thought experiment, Häggqvist points out that the argument from cross-world extensional difference to difference in content is not valid. Putnam, of course, is quite aware of this, and this may be an additional reason why he places the two planets in the same world.

¹² Frank Jackson stresses this point in a recent discussion of metaphysical necessity, and gives an explicit defence of the idea that it is part of the meaning of ‘water’ that it

names a natural kind: “Water is a rigid designator for the kind common to the watery exemplars we are ...acquainted with. This is what we grasp when we come to understand the word.” (1998, 49)

¹³ Soames (2002) argues that it is far from clear how the notion of rigidity can be applied in the case of general terms, such as ‘water’ and ‘heat’, since these function as predicates. This introduces a complication that I will not be able to address here. It is of interest to note, however, that Soames nevertheless believes that Kripke is right to suggest that natural kind terms are similar to names, in that they share a certain semantic property with names – i.e. the property of being non-descriptive (Soames 2002, 265). This property, Soames argues, is more fundamental than that of rigidity. The talk of semantic properties, suggests that Soames would grant that it is intrinsic to the semantics of a term such as ‘water’ that it is a natural kind term. At the same time, however, Soames wishes to allow for the possibility that a purported natural kind term fails to name a natural kind, without reference failure. Although he does not commit to a particular view on how this is to be reconciled with Kripke’s view, Soames argues that “the resolution of this issue is potentially relevant to the account of the actual semantic contents of natural kind terms like *water*...” (ibid. 284).

¹⁴ The disjunctivist solution has been suggested to me at several occasions when presenting earlier versions of this paper.

¹⁵ I discuss this further in my “Self-Knowledge, Externalism and Incomplete Understanding” (unpublished MS). For a critical discussion of McLaughlin and Tye see also McKinsey 2002.

¹⁶ Paul Boghossian argues against a view of McLaughlin and Tye’s sort on related grounds. Boghossian objects that this view implies that it cannot be known a priori whether *water* is an atomic concept or a compositional one, and that this is

metaphysically dubious. It is hard to see, he writes, “how the compositionality of a concept could be a function of its external circumstances in this way” (1997, 172).

Similarly, I want to suggest, it is hard to see how the rigidity of a term could depend on external circumstances in this way.

¹⁷ Kripke 1972, p.21, fn 21. Kripke distinguishes ‘de jure’ rigidity from ‘de facto’ rigidity where a description happens to designate rigidly since it uses a predicate that is true of the same object in every possible world (an example of the latter, Kripke suggests, is “the smallest prime”). Since proper names and natural kind terms are not de facto rigid Kripke suggests that what he is concerned with is ‘de jure’ rigidity.

¹⁸ For a full defense of descriptivism see Häggqvist and Wikforss, “The Semantics of Natural Kind Terms” (unpublished MS).

¹⁹ It is simply a mistake to assume that descriptivism must limit itself to superficial properties. David Lewis makes this point in a discussion of Twin Earth. Arguing that water is a cluster concept, he suggests that among the conditions in the cluster is that water is a natural kind, and that this is something traditional versions of the cluster theory failed to notice (Lewis 1994, 424).

²⁰ This is not to suggest that there is a simple formula telling us when to revise the belief that something is a natural kind. In Putnam’s thought experiment, twin water is said to share all of the macrolevel qualities of our water, and yet have a different microstructure, and so the existence of twin-water constitutes a straight-forward falsification of the assumption that ‘water’ names a natural kind. In the real world, things are messier and a great deal of indeterminacy is to be expected. Microlevel differences may lead only to small macrolevel differences, and so we are left with a choice as to how we should revise our beliefs. See LaPorte 1996 for some nice examples of this.

²¹ Kent Bach suggests a related point: “That water is a natural kind and mud is not does not establish a difference in semantic category between ‘water’ and ‘mud’.” (Bach 1987, 290).