

This is an accepted manuscript of an article published by Taylor & Francis in *Australasian Journal of Philosophy* on 21 May 2015, available online:
<http://www.tandfonline.com/doi/full/10.1080/00048402.2015.1046384>

NATURAL KINDS OF SUBSTANCE

Stephen Law

Abstract

This paper presents an extension of Putnam's account of how substance terms such as 'water' and 'gold' function and of how a posteriori necessary truths concerning the underlying microstructures of such kinds may be derived. The paper has three aims:

- (1) to refute a familiar criticism of Putnam's account: that it presupposes what Salmon calls an 'irredeemably metaphysical, and philosophically controversial, theory of essentialism'. I show how all the details of Putnam's account – including those Salmon believes smuggle in such essentialist commitments – can be squared with a rejection of any such essentialist metaphysics.
- (2) to reveal why Steward is wrong to suppose that, by helping himself to the claim that 'H₂O' is a rigid designator of a substance, Kripke, too, presupposes something controversially 'metaphysical'.
- (3) to show how my proposed account also sidesteps a variety of objections raised by Needham and others who argue that Kripke's and Putnam's accounts of how 'water' and 'gold' function founder upon the sheer microstructural complexity of the phenomena in question.

Keywords: Putnam, essentialism, natural kind.

1. Putnam and Donnellan on How 'Water' Functions

On Putnam's view, we discovered empirically that necessarily, something is a sample of water only if it is a sample of H₂O. According to Putnam, the possibility of our making such a discovery arises because of how 'water' is defined. We define the term by reference to samples of stuff in our environment, in such a way that an underlying and initially unknown microstructural property of that stuff is allowed to play a role in determining whether something is a sample of 'water', thus defined.

Allowing unknown properties to play a role in determining the conditions under which something qualifies as a so-and-so is not difficult. Take Donnellan's example (in an unpublished paper referenced by Salmon [1982: 165]) of 'watchamaycallit' stipulatively introduced thus:

Necessarily: someone is a *watchamaycallit* if and only if they have the same marital status as that currently possessed by *him* (pointing).

Suppose we subsequently discover that the person indicated is unmarried. Then, given that if two people share the same marital status (within or across possible worlds), then one is unmarried if and only if the other is, we're now in a position to deduce that that:

Necessarily: someone is a *watchamaycallit* iff they are unmarried.

This last necessity is a posteriori. In order to know that it obtains, we need to know the marital status of the person indicated, and that is knowledge that can only be acquired empirically. Yet clearly, despite being a posteriori, the final necessity is not, as it were, built

into the world. All the world contributes towards the truth of the conclusion is a matter of fact – that the person indicated is unmarried. The modal status of the conclusion traces back to our definition. The derived necessity may be a posteriori, but it is clearly stipulative in origin.

Putnam supposes ‘water’ is defined in a somewhat similar way. He says that the extension of a term like ‘water’ in part ‘depends upon the actual nature of the particular things that serve as paradigms, and this nature is not, in general, fully known to the speaker. Traditional semantic theory leaves out...the contribution of the real world’ [1973: 711]. In defining ‘watchamaycallit’ we use the indexical ‘him’ so as to allow an initially unknown property of the person to whom we refer to play a role in determining under what circumstances someone qualifies as a watchamaycallit. On Putnam’s view, we allow a similar ‘real world’ contribution when we give an ostensive definition of ‘water’. For example, I might point to the liquid in my glass and say:

- (1) Necessarily: something is a sample of water iff it is a sample of *dthat*¹ (the same substance of which *this* is a liquid sample)

Suppose we subsequently discover:

- (2) This (liquid sample) has the chemical composition H₂O.

Given that:

¹ ‘Dthat’ is Kaplan’s indexical operator that transforms any given singular term into a rigid designator.

(3) Being a sample of the same substance consists in having the same chemical composition

it then follows that:

(C) Necessarily, every sample of water has the chemical composition H₂O

This final necessity is a posteriori, according to Putnam. In order to know the conclusion, we need to know the truth of the second premise, and that can only be established empirically.

Call the above argument *the water argument*.

There's at least one noteworthy difference between the watchamaycallit and water arguments, a difference pointed out by Donnellan (quoted by Salmon [1982: 166]):

'Water is H₂O', assuming the correctness of Putnam's account, is doubly exotic.

Involved in our knowledge of its truth, even though it is necessarily true, is not only the empirical discovery that the stuff we call water has the chemical structure H₂O, but also the knowledge, if that is the right term, that we have, and the ancients did not have, that two liquids are the same only if they have the same chemical structure. And that also is not a priori knowledge.

Why does Donnellan suppose it is not a priori knowledge that two liquids are the same only if they share the same chemical structure? Certainly, unlike the principle that if two people share the same marital status then one is unmarried if and only if the other is, it isn't a trivial conceptual truth. If it were, then it could have been known even back in 1750. Yet various

scientific developments had to take place before this condition on being same-substance related could be known.

2. The Alleged Background Metaphysics

We now arrive at Salmon's suggestion that the third premise of the water argument appears to involve a commitment to 'an irrevocably metaphysical, and philosophically controversial, theory of essentialism' [1982: 217]. The third premise of Putnam's argument may be no trivial conceptual truth. But neither does it restrict itself to stating how things actually are. As Salmon explains, if Putnam's conclusion is to follow, premise (3) has to make a very strong modal claim: a claim concerning not just how things stand with respect to the actual world, but with respect every possible world. It must say something like:

(S3) A sample x in possible world w_1 is a sample of the same substance as sample y in possible world w_2 only if x in w_1 possesses the same chemical composition as y possesses in w_2 .²

Unless the third premise places such a constraint on what it is for samples to be same substance related within and across worlds, Putnam's conclusion that any possible sample of water has the chemical composition H_2O won't follow. So the water argument appears to depend crucially on a premise that, when unpacked, turns out to involve a contentious, non-trivial modal claim. Indeed, Salmon insists that (S3) is established, if at all, not by science, but by 'metaphysical' thought experiment [1982: 185-6]:

² My unpacking of premise (3) is a simplified version of Salmon's unpacking. They differ slightly, but not in a way that undermines points made here. See Salmon [1982: 179-80].

Unmasked, [(3)] looks instead like a principle whose truth is established in good part by means of conceptual analysis and the reflective methods of the metaphysician. One tests [it] not by laboratory experiment, but by thought experiment. One might reason, for instance, as follows: suppose that there were a substance whose chemical structure is different from the actual structure of a given substance, say, water. Could this possible substance be the very same substance, water, nevertheless? No. It must be a different substance, namely, whatever substance has that particular chemical structure. This line of thought is an activity neither of science nor of the philosophy of language. It is metaphysics and nothing else.

While Salmon finds this background metaphysical theory plausible, he insists it is not established by the water argument. Rather, it's *presupposed* by the water argument. It's smuggled in via that third premise.

One reason the discovery of this hidden metaphysical commitment is important, thinks Salmon, is that the water argument is intended by Putnam to derive a significant non-trivial essentialist conclusion from nothing more than the theory of direct reference in combination with a philosophically uncontentious empirical discovery. True, Putnam also argues for the claim that necessarily: something is a sample of water only if it has the chemical composition H_2O using his twin-earth thought experiment. But thought experiments rely on appeals to philosophical intuition, and intuition is a fickle thing. As Salmon sees it, the water argument is supposed to by-pass any such appeal to intuition. It is intended to deliver Putnam's conclusion using nothing more than the theory of direct reference in

combination with a straightforward empirical discovery. Salmon believes he has shown the water argument fails in this respect because he has shown that it depends on a controversial, non-trivial essentialist thesis established, if at all, by thought experiment.

In fact, we might suspect that Salmon's unpacking of (3) reveals that Putnam's account of how natural kind terms like 'water' are defined does, after all, presuppose that the world comes with exotic modal facts 'built in'. When I point to the stuff in my glass and define 'water' in (1) above, I do so in such a way that what Putnam calls 'the contribution of the real world' is two-fold. I allow nature to contribute, not only the philosophically uncontroversial fact that sample indicated has the chemical composition H_2O , *but also the fact that samples are same substance related within and across worlds only if they share the same underlying chemical composition*. Indeed, Putnam's account of how 'water' functions might seem to require the existence of a kind of modal grid embedded in nature at the level of microstructure: a grid delineating kinds for which we then only need introduce labels. Call such kinds Natural Kinds (capital 'N' and 'K').

Now let's return to the related issue of 'non-trivial essentialism'. Salmon believes that Putnam and Donnellan are attempting to derive a 'non-trivial' form of essentialism from nothing more the theory of direct reference plus a philosophically uncontroversial scientific discovery. Salmon believes he has shown their attempt fails. But what does Salmon mean by 'non-trivial essentialism'? He explains largely by examples. In an earlier paper Salmon contrasts 'non-trivial' essentialism with forms of essentialism that attribute trivially essential properties to things, giving as an example the property of *not being both red and not red* [1979: 703]. In *Reference and Essence*, Salmon again explains by means of examples. To claim Hesperus has the essential property of *being identical with Phosphorus*, suggests Salmon, is to commit oneself only to a trivial form of essentialism, despite the claim being a

posteriori. Salmon says that if this, if a brand of essentialism at all, is a brand 'of the most trivial and innocuous kind' [1982: 83]. That's because the necessity traces back to the trivial truth that every object has, as a matter of necessity, the property of being the very object it is, that being little more than a 'boring truism'. By contrast, Salmon says the claims that it is an essential feature of this table that it was originally made from a particular hunk of wood and that it is an essential property of the kind liquid (pure) water that it is composed of two parts hydrogen and one part oxygen both involve essentialism of a 'non-trivial' sort [1982: 84-7].

Now if premise (3) of the water argument requires that there be Natural Kinds, then the project in which Salmon supposes Putnam and Donnellan are engaged is indeed a failure. A commitment to Natural Kinds is surely a commitment to what Salmon calls 'an irrevocably metaphysical, and philosophically controversial, theory of essentialism' [1982: 217]. However, Salmon doesn't go as far as to suggest Putnam helps himself to Natural Kinds.

Having separated out the distinct though nevertheless related issues of whether Putnam's account presupposes (i) Natural Kinds and/or (ii) and 'non-trivial essentialism', I will now explain why Putnam's account may presuppose neither.

3. The Wizzle Argument

Suppose a magician welcomes us into a warehouse over which, he tells us, he has cast a magic spell. The warehouse is full of boxes. Each box comes in one of several different colours. All of the boxes are currently locked. The magician tells us that his spell makes the outward colour of the boxes causally dependent on the combination of objects they contain. There is a certain object/feature combination possessed by the boxes, the precise variety of which determines the outward colour of the box. For example, says the magician, it might be that the boxes contain wooden shapes, with a box containing two cubes and one sphere

caused to be red, a box containing three spheres caused to be puce, and so on. Or it might be the combination of coloured marbles contained by the boxes that determines box colour, or perhaps the combination of differently sized beanbags. The magician refuses to provide any further details about what the causally relevant object/feature combination is. He also informs us that a small percentage of the boxes are rogue – they are exceptions to these rules he has introduced.

Believing our magician to be trustworthy, we suppose there is likely to be a hidden similarity between (most of) the red boxes. We suppose they're probably alike not just in terms of their outward colour, but also in terms of their contents. Ditto the green boxes. We also suppose that there is a hidden difference – a difference in terms of the combination of objects they contain – between the red boxes and green boxes etc.

Now suppose we introduce some terms in such a way as to latch onto these hypothesized, currently unknown, hidden similarities and differences between the boxes. First, we stipulatively introduce the relational term 'same sort of woozle' like so (pointing to a pile of boxes of many different of colours):

(SW1) x in possible world w_1 is an example of the *same sort of woozle* as y in possible world w_2 iff x in w_1 is a box and y in w_2 is a box and x in w_1 contains the same variety of *dthat* (the inner object/feature combination of *these* boxes that is (i) the same in most of the same-coloured boxes, (ii) different in most of the differently-coloured boxes and (iii) usually determines the box's outward colour) as y contains in w_2 .

We then stipulatively introduce the common noun ‘wizzle’³ like so (pointing at a certain red box, B):

(1*). Necessarily: something is a *wizzle* iff it is an example of *dthat* (the same sort of woozle of which *this* is currently an example).

We similarly define other terms, e.g. we define ‘wazzle’ by reference to a puce box, ‘wozzle’ by reference to an orange box, and so on.

Later, the magician unlocks the boxes and allows us to look inside. We make two empirical discoveries. First we discover:

(SW2) The inner object/feature combination that is (i) the same in most same-coloured boxes, (ii) different in most differently coloured boxes and (iii) usually determines outward box colour, is marbles/colour.

That’s to say, it’s the combination of coloured marbles a box contains that, under the magician’s spell, usually causally determines the box’s outward colour. Call the combination of coloured marbles a box contains its *marble composition*. We also discover:

(2*). This (box B) contains just two red and one yellow marbles (has, let’s say, the *marble composition* R₂Y)

(SW1) and (SW2) allow us to deduce:

³ The terms ‘wizzle’ and ‘woozle’ are borrowed from A.A. Milne’s *Winnie-the-Pooh*.

(3*). Being the same sort of woozle as something consists in being a box with the same marble composition (a box x in possible world w1 is an example of the same sort of woozle as box y in w2 iff x has the same marble composition in w1 as y has in w2)

From (1*), (2*) and (3*) we can then make the further deduction:

(C*): Necessarily: Something is a wizzle iff it is a box with the marble composition
R₂Y

Call this *the wizzle argument*.

Notice that, given the conclusion of the wizzle argument, were you to be presented with one of the rogue red boxes (one of the boxes that are exceptions to the magician's spell) prior to the discoveries about box contents and you concluded that it was a wizzle, you would be mistaken. The box would meet the outward criteria on the basis of which we would at that point judge something to be a wizzle. But, crucially, it would *lack the right marble composition*. It would be a *fool's wizzle*.

Notice that the contribution of the 'real world' so far as the truth of the conclusion of the wizzle argument is concerned is here restricted to two philosophically uncontroversial, and empirically discoverable matters of fact: the fact that box B contains two red and one yellow marble, and the fact that marble composition is what causally determines box colour.

Finally, regarding the stipulative introduction of the terms 'wizzle' and 'same woozle', notice both involve certain empirical presuppositions. This was also true of Donnellan's definition of watchamaycallit. When Donnellan defines 'watchamaycallit', he

supposes that he is pointing at someone with a marital status. But perhaps he is pointing at a mannequin. Then his definition fails to latch on to any marital status. Similarly, in introducing 'same woozle' by means of (SW1), we are presupposing that there is, indeed, some inner object/feature combination of the indicated boxes that meets (SW1)'s conditions (i)-(iii). If the magician lied to us and there's no such inner combination, our definition will also fail.

4. A Sketch of How 'Same Substance' May Function

The analogy I'm drawing between the wizzle and water arguments should be obvious enough. Premises (1*), (2*), (3*) and (C*) of the wizzle argument are intended to parallel (closely, if not perfectly) premises (1), (2), (3) and (C) of Putnam's water argument. The key difference between the two arguments is that there's a *further tier of argument involved in the latter: that by which (3*) is derived from (SW1) and (SW2)*. My suggestion, of course, is that the expression 'same sort of substance' may be defined in something like the way 'same sort of woozle' is defined, and that, consequently, premise (3) of Putnam's water argument may be derived in much the same way as premise (3*) of wizzle argument.

How might such a definition of 'same substance' go? Here's one suggestion. Suppose that, before discovering atomic and molecular constitutions, we suspect that amongst the various microstructural features of various samples of stuff in our local environment is at least one - perhaps several - underlying levels of constitution that play a significant role in determining those properties on the basis of which we have sorted various samples into groups. Call such group-sorting observational properties the P properties of those samples/groups (e.g. one group's P properties might be: clarity, liquidity, lack of odour, thirst-quenching character, and 100°C boiling point). We suppose that samples within groups

will, in most cases, share these levels of constitution and that samples from differing groups will mostly differ in at least one of these levels of constitution (Note that, even back in the mid 17th Century, Locke suspected such underlying constitutions existed, and warned his contemporaries against attempting to annex terms like ‘water’ to them).

Now suppose that, based on this empirical supposition, we point to a large number of samples ordered into sample groups and define ‘same sort of substance’ stipulatively like so:

(SS) a sample x in possible world w1 is a sample of the *same sort of substance* as sample y in possible world w2 iff (for each of the below-mentioned levels of microstructural constitution that they share⁴) x in w1 possesses the same varieties of *dthat* (those levels of microstructural constitution of *these* samples that, where present in a given sample group, (i) is the same in most same-grouped samples, (ii) differs in most differently-grouped samples, and (iii) plays a significant role in causally determining most of the P properties of the various samples) as y in w2.

Note the obvious intended similarity between our stipulative definitions of ‘same woozle’ (SW1) and ‘same substance’ (SS).

Suppose we subsequently discover that those levels of microstructural constitution that meet conditions (i)-(iii) in (SS) are atomic structure (or more specifically, atomic number – number of protons per atom) and chemical composition (or, more accurately, molecular structure – see below). More specifically, we discover that it’s the atomic number(s) of the constituent atoms and (where present) the molecular structures they go to make up that is the

⁴ We might add ‘of which they must share at least one’. Otherwise it will be left open whether samples lacking molecular and atomic compositions (XYZ perhaps?) might nevertheless qualify as ‘water’. Further refinements may also be required.

same in most same grouped samples, differs in most different grouped samples, and plays a significant role in determining P properties. We might then deduce Salmon's (S3). If we have also previously also defined various kind terms such as 'water', 'gold', 'iron', 'alcohol', and so on by means of stipulative definitions such as (S1) that *utilize the expression 'same substance' as defined by (SS) above*, we might then further deduce e.g. that necessarily, something is a sample of water only if it is a sample possessing the chemical composition H₂O.

This is merely one suggestion concerning how 'same substance' might be defined in a manner roughly analogous to the way in which 'same woozle' is defined above. Variants and refinements are possible.

Of course, few if any of us have ever *explicitly* stipulated that 'same substance' is to function in the manner set out above. But perhaps (SS) more or less accurately captures our linguistic intentions regarding 'same substance'. It does not strike me as outrageous to suppose that, say, many of Locke's scientific contemporaries intended that 'same substance' should function in something like that way. Still, let me be clear that my proposal here is that *if* speakers have such linguistic intentions, then Putnam's conclusion might be derived in the way I describe. Whether speakers do have such intentions is an empirical question.

Experimental studies into how speakers intend natural kind terms to function have produced mixed results. While some studies suggest speakers do broadly intend natural kind terms to function in a Putnamian way (see Jylkkä, Railo, and Haukioja [2009]), others have drawn more qualified or skeptical conclusions (see for example Hampton, Estes, and Simmons [2007] and Braisby, Franks, and Hampton [1996]). There is also a question mark over whether our historical practice has been to use substance terms with the intention that they should function in the way Putnam supposes (for example, LaPorte [2004] provides some

apparent counter-examples to that claim). Whether the general population or some scientific sub-community have, or previously had, the intention of using both ‘water’ and ‘same substance’ in the suggested manner is an issue I put to one side here, and so I leave it an open question whether terms like ‘water and ‘same substance’ *do* in fact function in the suggested manner.

Finally, a note regarding *definitional success*: just like Donnellan’s definition of ‘watchamaycallit’, our definition (SS) involves *an empirical presupposition that may not be met*. Perhaps there exist no underlying microstructural constitutions of the sort posited by (SS). This possibility, and related objections, are discussed towards the end of this paper.

5. Some Conclusions

We can now draw five significant conclusions:

1. Salmon takes the fact that premise (3) of the water argument, when unpacked, turns out to be neither a trivial conceptual truth, nor restricted to reporting some mere matter of fact, to indicate that it involves a commitment to ‘an irredeemably metaphysical, and philosophically controversial, theory of essentialism’. But notice that *premise (3*) of the wizzle argument shares these same features*. It too is neither a trivial conceptual truth nor restricted to reporting some mere matter of fact. It too makes a claim concerning all possible worlds. And yet (3*), I take it, involves no commitment to anything irredeemably metaphysical. But then there’s no good reason to conclude, as yet, that Putnam’s water argument involves any such commitment. ‘Same sort of substance’ might be defined by means of something like (SS). In which case, (3) might yet be delivered in much the way (3*) is, by means of a deductive

argument combining nothing more than a stipulative definition and a metaphysically innocuous discovery.

To digress for a moment (and draw a contrast with a rival account of how ‘water’ functions): I’m not the first to suggest that the necessity that something is a sample of water only if it has the chemical composition H_2O traces back to necessities that are trivial, because stipulative, conventional, and/or analytic in nature. The most worked out account to date is probably Sidelle’s, who suggests [1989: 37] that it is a

matter of convention, say, that if something is a chemical kind, then it has its chemical microstructure necessarily. Thus, if we can add the ostensibly nonmodal ‘Water is a chemical kind’ and ‘The microstructure of water is (actually) H_2O ’. we will be able to derive that it is necessary that water is H_2O , and all the modal force of this conclusion will be derived from our general principle, which we are supposing to be analytic.

In attempting to show how the necessity regarding water’s microstructure can be traced back to linguistic conventions, Sidelle shares an aim of mine: to reveal how the a posteriori necessity that something is a sample of water only if it is H_2O might be derived without presupposing any controversial essentialist metaphysics.

However, Sidelle traces that necessity back through just a *single* tier of argument to a *single* supposedly analytic truth: that chemical kinds have their chemical microstructures necessarily. But then Sidelle’s account of how ‘water’ functions fails, as it stands, to acknowledge or account for what Donnellan calls the ‘doubly exotic’ status of the necessity that something is a sample of water only

if it is H₂O. What Donnellan notes, and Salmon focuses attention on, is the fact that the third premise of Putnam's water argument appears to be no trivial, analytic truth. If it really is a condition on samples being same substance related within and across world that they share chemical composition, and if this is indeed no mere conventional, analytic truth (which, like Salmon, I find intuitively plausible), then this requires some explanation given Sidelle's view that all necessity is conventional in origin. Sidelle's account fails in this respect. It is that explanation that my suggested account supplies. Unlike Sidelle's single-tiered account, my two-tiered account does justice to all the details of Putnam's account of how 'water' functions, *including those details that Salmon supposes smuggle in 'non-trivial' essentialist commitments*. I note Haukioja [forthcoming] also offers an interesting proposal as to how Putnam's conclusion that necessarily: every sample of water has the chemical structure H₂O, might be derived without making any non-trivial essentialist assumptions. Haukioja's strategy, like Sidelle's, also involves dispensing with Putnam's third premise, and so does not directly address, as I do here, the question of its alleged 'metaphysical' status. My aim here is to show, not just how Putnam's conclusion might be derived without presupposing anything contentiously 'metaphysical', but how it might be derived *in just the way Putnam suggests*, via his third premise.

2. Salmon suggests premise (3) of the water argument is established, if at all, by thought experiment. Notice that premise (3*) of the wizzle argument *can* be tested by thought experiment. Having established that the causally relevant object/feature combination is marble composition, we might test our intuitions with the following question: 'Consider a

box whose marble composition is different from the actual structure of a given woozle, say, a wizzle. Could this possible woozle be a wizzle, nevertheless?' However, while we could appeal to thought experiment to test (3*), we don't need to. That's because (3*) can just be derived directly from (SW1) (a definition) and (SW2) (a metaphysically innocuous empirical discovery). And if (3*) might be established by such means, then perhaps so might (3). There's no reason yet to think (3) can *only* be established by means of thought experiment, i.e. that it cannot be derived in the same manner as (3*).

3. As we have seen, Salmon is concerned with whether 'non-trivial' forms of essentialism can be derived from the theory of direct reference plus certain philosophical uncontroversial facts (such as the fact that this liquid does *in fact* have the chemical composition H₂O). Salmon considers the water argument an attempt to achieve this. He may be right that the water argument, as it stands, fails in this regard, but, as I've explained, Salmon overlooks the possibility of introducing a second tier of argument by which (3) is derived in much the same way as (3*). However, supposing the conclusion of the water argument (C) *could* then be derived in much the way as the conclusion of the wizzle argument (C*), should we conclude that we *would* have succeeded in deriving a 'non-trivial' form of essentialism from the theory of direct reference plus some metaphysically innocuous scientific discoveries? Surely not. For why suppose that (C) involves any more commitment to a 'non-trivial' form of essentialism than does (C*)?

4. Note that what Donnellan terms the 'doubly exotic' status of (C) is mirrored by (C*). Both conclusions appear to involve, as it were, a double dose of a posteriority. Both kinds are carved out by constraints on sameness that can only be discovered empirically. However,

wizzles are not a natural kind. It would be a mistake, then, to conclude that what marks out natural kinds is the fact that conclusions concerning their essential properties involve such a double dose of a posteriority.

5. Finally, I note the wizzle argument has significant implications for an argument of Kripke's concerning water:

If 'a' and 'b' are rigid designators, and $a = b$, then necessarily: $a = b$

'Water' and 'H₂O' are rigid designators

Water = H₂O

Necessarily: water = H₂O

Kripke's argument, like the water argument, also delivers a conclusion that is a posteriori, (because the third premise can only be established empirically). However, the conclusion is stronger than that of the water argument.

Interestingly, Kripke's argument has also been accused of smuggling in something contentiously 'metaphysical'. Steward [1990] suggests that, by helping himself to the claim that 'H₂O' rigidly designates a substance, Kripke helps himself to a philosophically controversial thesis. For 'H₂O' would appear to abbreviate a description – something like 'the chemical kind composed of molecules made up of one atom of oxygen and two of hydrogen'. If this description does rigidly designate a substance, says Steward, it doesn't do so *de jure*. The rigidity of 'H₂O' is not like the rigidity that trivially attaches to an ordinary proper name. The rigidity of 'H₂O' with respect to a substance will be *de facto*: it requires the truth of what Steward [1990: 396] calls a 'substantive metaphysical presupposition', namely, that

‘chemical composition is, as a matter of metaphysical fact, the basis of the same substance relation’ [1990: 393]. According to Steward, this presupposition is established, if at all, by what Steward calls ‘metaphysical considerations’ – i.e. by thought experiment.

Soames [2002: 273-310], on the other hand, maintains Kripke’s argument requires no such metaphysical background theory. According to Soames, the rigidity of ‘H₂O’ with respect to a substance requires only the principle that samples are same substance related within and across worlds if and only if they share what he calls the same ‘basic physical constitution’. And *this* principle, suggests Soames, is no deep metaphysical truth. Rather, *it’s analytic*: it’s a straightforward consequence of how ‘substance’ is defined (for Soames supposes ‘substance’ is explicitly defined as a basic physical constitution). So, concludes Soames, Kripke’s conclusion can be derived without appeal to any substantive metaphysical presuppositions.

Salmon rejects Soames’s suggestion. Soames’ principle, suggests Salmon, is no trivial, analytic truth. According to Salmon, in order for it to pick out the relevant underlying microstructural features (atomic or chemical compositions), what Soames means by ‘basic physical constitution’ must be defined to mean atomic or chemical composition (i.e. so that we can then know a priori that ‘substances’, thus defined, are chemical elements or compounds). But then, concludes Salmon [2003: 489]: ‘...this makes the notion of substance reliant to a large degree on modern chemistry, and hence far more theory-laden than the original introduction of the word ‘water’ could have been, given the general scientific ignorance of the day.’

I draw three morals concerning this debate between Steward, Salmon, and Soames.

The first moral is: even supposing Salmon is right that Soames' principle, when unpacked, is no trivial, analytic truth, but rather something like (S3)⁵, that's not yet to say that there need be anything interestingly 'metaphysical' about either it or the rigidity of 'H₂O'. To see why, compare 'R₂Y'. Notice 'R₂Y' is *also* a rigid designator: it designates with respect to any possible world the same sort of woozle. Why so? Well, (3*) says that boxes are same woozle related within and across worlds if and only if they share the same marble composition. Thus, (3*) entails that there is no possible world with respect to which 'R₂Y' designates anything other than the sort of woozle it designates with respect to the actual world. Now (3*) is no trivial, analytic truth. Yet there's nothing interestingly metaphysical about either it or the consequence that 'R₂Y' is de facto rigid with respect to a certain sort of woozle. But then neither need there be anything interestingly metaphysical about the non-analytic (S3) or the consequence that 'H₂O' is rigid with respect to a certain sort of substance.

Secondly, Steward shouldn't assume that if the rigidity of 'H₂O' with respect to a substance is de facto, then that rigidity will be established, if at all, by way of metaphysical thought experiment. As we've seen, the rigidity of 'H₂O' with respect to a substance might yet be derived in much the same way that the de facto rigidity of 'R₂Y' with respect to a kind of woozle.

The third moral is this. Salmon insists that, in order to succeed, Soames' account requires, implausibly, that it be an analytic truth that samples are same substance related within and across worlds if and only if they share the same chemical composition. My suggested account, by contrast, by allowing that the expression 'same substance' may be

⁵ What's needed to deliver the conclusion that 'H₂O' is rigid with respect to a substance is the bi-conditional strengthening of (S3).

defined in such a way as to allow the world to make a contribution to determining the conditions under which it is correctly applied, actually neatly explains the non-analyticity of that principle. It also reveals why at least Soames's conclusion is correct: the derivation of Kripke's conclusions concerning the necessary chemical composition of water need not rely on any controversial metaphysical essentialist thesis.

6. Two Putnamian Accounts of How 'Water' Functions

We are now presented with two rather different accounts of how natural kind terms like 'water' function, both of which are consistent with both the details of Putnam's account and his method of deriving the conclusion that necessarily, every sample of water has the chemical composition H_2O . On the first 'metaphysical' picture, what Salmon calls an 'irredeemably metaphysical, and philosophically controversial' theory of essentialism is presupposed. Some might go further and suppose that, on Putnam's account, the definition of terms like 'water' presupposes that there are Natural Kinds. When we give ostensive definitions of natural kind terms like 'water', we do so in such a way that our terms can latch on to (what we at least take to be) what Putnam has elsewhere called 'objective essences' which are somehow 'out there' [1982: 157]. By contrast, on the second picture sketched out here, 'water' and 'same substance' function in something like the way 'wizzle' and 'same woozle' function. No philosophically controversial background theory of essentialism is presupposed.

7. The 'Empirical Presuppositions' Problem

Is my suggested second picture correct? Do 'water' and 'same substance' function in something like the way 'wizzle' and 'same woozle' function?

Arguably not. As noted above, perhaps we just don't intend 'same substance' to function in anything like the way 'same woozle' functions. Some might object, in particular, that my suggested account attributes rather more sophisticated and specific linguistic intentions and/or background knowledge to speakers than might plausibly be attributed to them. However, my account doesn't require speakers possess *very* much more sophisticated linguistic intentions than those attributed to speakers in the wizzle case. Certainly, no background knowledge of atomic and molecular structures, chemical compositions and the like is required in order to define 'same substance' in the manner suggested.

However, even if 'same substance' were defined in something like the way 'same woozle' is defined, there remains the objection that such a definition will fail because its empirical presuppositions are not met: the world just isn't microscopically ordered and causally structured in the way required for such a definition to succeed.

The suggestion that common nouns like 'water' and 'gold' function in a Putnamian way has been widely criticised on the grounds that the empirical presuppositions of such definitions are not met. Needham, for example, insists that 'whatever water's microstructure is, it is very complicated', far more so than Putnam and Kripke acknowledge [2011: 11]. For example, says Needham, 'water in the form of ice has a very different microstructure from water in the form of a liquid, and again a very different microstructure from water in the gas phase' [2011: 9]. But then there is no *single* underlying microstructure that 'water' picks out.

Needham offers several other criticisms of what he calls 'microessentialism'. He maintains that, contrary to what Putnam and Kripke seem to assume, 'H₂O' *isn't even a microdescription*. According to Needham, "'H₂O'" is a compositional formula, containing the information that the compound is composed of hydrogen and oxygen in the fixed gravimetric proportions of 1 : 8 converted into a scale of equivalents' [2011: 9]. To describe a substance

as 'H₂O', says Needham, is thus to describe it as 'being a compound of hydrogen and oxygen in the proportions 2 moles to 1' and this is 'entirely macroscopic information'. Needham concludes that 'anyone wanting to give a microdescription of water who simply offers "H₂O" fails miserably' [2011: 9].

Other substances Needham supposes cause trouble for Putnam's account include isomers – molecules with the same chemical composition but different molecular structures (they contain the same number of atoms of each element, but in different structural combinations). As Needham points out, dimethyl ether and alcohol share the same *compositional* formula, C₂H₆O, yet are distinct substances with their own *structural* formulas - (CH₃)₂O and C₂H₅OH respectively. But then sameness of chemical composition is insufficient to qualify samples as being same substance related. [2011: 9]

Needham also notes that terms such as 'gold' pick out not chemical compounds but elements the underlying microstructural essences of which are supposedly atomic, not molecular, constitutions (atomic numbers, to be precise). The implication seems to be that what constitutes the microstructural essence of a substance kind is treated inconsistently by Kripke and Putnam across substance kinds.

Objections of this sort are considered by many to spell doom for any broadly Putnamian account of how substance terms function and how a posteriori necessary truths concerning microstructure may consequently be derived. Before nailing down the coffin lid on Putnamian 'microessentialism', however, it's worth noting one very significant advantage of my proposed Putnamian of how 'same substance' and associated substance terms may function: *it appears to deal with many of these objections, including all those presented by Needham above.*

Consider again the suggested definition of 'same substance':

(SS) a sample x in possible world w_1 is a sample of the *same sort of substance* as sample y in possible world w_2 iff (for each of the below-mentioned levels of microstructural constitution that they share) x in w_1 possesses the same varieties of *dthat* (those levels of microstructural constitution of *these* samples that, where present in a given sample group, (i) is the same in most same-grouped samples, (ii) differs in most differently-grouped samples, and (iii) plays a significant role in causally determining most of the P properties of the various samples) as y in w_2 .

Now first consider Needham's observation that water has a great many microstructural features – including different microstructural features accounting for its differing phases. Does this cause a problem for my suggested account? No. Given water is defined in the manner of (S1) and 'same substance' in the manner of (SS), microstructures specific to phases are unlikely to be included in the essences of substance kinds. For many of the sample groups by which 'same substance' is defined are likely to include multiple phases (e.g. solid and molten gold, liquid and solid water, etc.), and thus more than one phase-specific microstructural feature. In which case condition (i) of (SS) will rule out phase-specific microstructural features (note that phase-specific microstructural features are also unlikely to play a significant causal role in accounting for *most* P properties of most sample groups, and so are also likely to be filtered out by condition (iii)).

Next consider isomers and Needham's objection that if sameness of chemical composition qualifies samples as being same-substance related, then samples of alcohol and dimethyl ether should qualify as belonging to the same substance, which they're not. Again, it's unclear why isomers present a difficulty for the suggested account. Notice, first, that

sameness of what Needham calls ‘chemical composition’ clearly won’t qualify samples as being same substance related given we define ‘same substance’ by means of (SS). (SS) is explicitly designed to pick out one or more microstructural constitutions, which, as Needham admits, his ‘chemical composition’ isn’t. Furthermore, note that isomers often differ markedly in terms of their observational properties. It is not just the number of different types of atom that go to make up a molecule of a substance that accounts for such observational properties of the sort likely to be included in P properties, but more specifically *the way in which they are combined structurally*. But then the depth and detail of microstructural constitution that best causally accounts for the P properties of our various samples (whether or not they happen to include isomers such as alcohol) is indeed likely to include *the kind of molecular structures captured by structural formulas*. Indeed, it is presumably such a specific molecular structure, rather than what Needham means by a ‘chemical composition’, that Kripke and Putnam really have in mind when they talk about ‘H₂O’ (this would, at least, be the more charitable reading). But then, if a specific molecular structure is a necessary condition for something to qualify as sample of alcohol, dimethyl ether won’t qualify. Thus Needham’s isomer objection is also dealt with.

Perhaps a defender of Needham will say that, since water has no isomers, and since (as understood by Needham) ‘H₂O’ *isn’t even a microdescription*, why insist on a microdescriptive characterization of water? Why not just say that water is H₂O, ‘H₂O’ being understood in Needham’s supposedly macroscopic way as ‘being a compound of hydrogen and oxygen in the proportions 2 moles to 1’? But this is to miss the point. Putnam’s concern is with how ‘water’ *was and is* defined, not with how we might now choose to (re)define it by means of some co-extensional description. For all Needham has said,

Putnam's account of how 'water' was and is defined (by reference to an unknown microconstitution subsequently discovered a posteriori) may still be correct.

What of Needham's observation that Putnam and Kripke switch between talking about molecular constitution when referring to the underlying essence of water and atomic constitution (atomic number) when referring to the underlying essence of gold? Doesn't this lack of consistency undermine any unified microessentialist account of what it is for samples to be same substance related?

It's not clear there is any lack of consistency here. Note, again, that (SS) explicitly allows that *more than one* level of microstructural constitution may be involved in determining whether or not samples are same-substance related. More than one level of constitution may meet conditions (i)-(iii). Thus sameness at more than one level of microstructural constitution may be necessary for samples to be same-substance related. And that sameness at the level of both the atomic numbers of constituent atoms *and* (where present) molecular structure is indeed required for sameness of substance *would then neatly explain why dioxygen (O₂), ozone (O₃), and other allotropes of the element oxygen (atomic number 8) are nevertheless considered distinct substance kinds by chemists*. So why, then, doesn't gold also have an essential molecular structure? In fact, gold doesn't have a molecular structure at all - solid gold is just the atoms organized into a lattice structure. (SS) explicitly allows that where a relevant level of microstructural constitution is missing, requirements concerning sameness at that level do not apply.

In short: by supposing that 'same substance' functions in the way 'same wozle' functions, we can neatly answer a variety of otherwise difficult questions posed by Needham and others that Putnam's original account left unanswered: questions about how ostensive definitions of terms like 'water' and 'gold' by means of definitional samples are able

successfully to lock on to certain specific microstructural features of those samples among the great many different microstructural features that they possess.

Needham and I are probably in agreement that we should resist the claim that there are Natural Kinds and that the world comes with modal facts concerning microstructure ‘built in’. No doubt *that* sort of essentialism should be rejected. But, unlike Needham, I see no reason as yet to reject either Putnam’s claims about how ‘water’ functions, or his claim that it’s an a posteriori necessary truth that something is water only if it has the chemical structure H₂O. Needham has as yet failed to show that either of these claims is mistaken.

8. Conclusion

This paper sketches a unified account of how natural kind substance terms may function. The account is an extension of Putnam’s original account. At the heart of my account lies the suggestion that the relational expression ‘same substance’ may *also* be defined so as to allow the world to make a contribution to determining the conditions under which the expression is correctly applied. (SS) illustrates how such a definition might go. Once the relational expression ‘same substance’ is introduced in this way, substance names may be defined by means of it in the manner Putnam suggests, and necessary a posteriori conclusions concerning the microstructural features might subsequently be deduced. My suggested development of Putnam’s account offers a number of advantages over his original, including:

1. The account neatly explains the correctness of the modal intuition Salmon and others have regarding the same substance relation – that it is a condition of samples being same substance related across worlds that they share the same (as he puts it) ‘chemical composition’.

2. The account explains how Putnam's a posteriori necessary conclusion regarding water might be derived from stipulative ostensive definitions and philosophically uncontroversial empirical premises *alone*, without making any significant metaphysical assumptions.
3. The account also delivers the thesis that 'H₂O' is a rigid designator, and again does so without making any significant metaphysical assumptions.
4. The account also has the advantage of sidestepping various objections to Putnam's account of the sort that have been raised by Needham – objections based on the sheer microstructural complexity exhibited by substance kinds.⁶

Heythrop College, University of London.

REFERENCES

Braisby, Nick, Franks, Bradley, and Hampton, James. 1996. Essentialism, Word Use, and Concepts, *Cognition* 59: 247-274.

Haukioja, Jussi forthcoming. On Deriving Essentialism From the Theory of Reference, *Philosophical Studies*.

Hampton, James A., Estes, Zachary, and Simmons, Sabrina 2007. Metamorphosis: Essence, Appearance, and Behaviour in the Categorization of Natural Kinds, *Memory and Cognition* 35/7: 1785-1800.

Jylkkä, Jussi, Railo, Henry, and Haukioja, Jussi 2009. Psychological Essentialism and Semantic Externalism: Evidence for Externalism in Lay Speaker's Language Use, *Philosophical Psychology* 22/1: 37-60.

LaPorte, Joseph 2004, *Natural Kinds and Conceptual Change*, Cambridge: Cambridge

⁶ I am grateful to Helen Beebe and an anonymous reviewer for helpful comments on earlier drafts.

University Press.

Needham, Paul 2011. Microessentialism: What is the Argument? *Noûs* 45/1: 1-21.

Putnam, Hilary 1973. Meaning and Reference, *Journal of Philosophy* 70/19: 699-711.

Putnam, Hilary 1982. Why There Isn't a Ready-Made World, *Synthese* 51/2: 141-167.

Salmon, Nathan 1979. How *Not* to Derive Essentialism from the Theory of Direct Reference, *Journal of Philosophy* 76/12: 703-725.

Salmon, Nathan 1982. *Reference and Essence*, Oxford: Basil Blackwell.

Salmon, Nathan 2003. Naming, Necessity, and Beyond, *Mind* 112/447: 475-492.

Sidelle, Alan 1989. *Necessity, Essence, and Individuation: A Defence of Conventionalism*, Ithaca N.Y.: Cornell University Press.

Soames, Scott 2002. *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity*, Oxford: Oxford University Press.

Steward, Helen 1990. Identity Statements and the Necessary A Posteriori, *Journal of Philosophy* 87/8: 385-398.