NATURAL KINDS AND MODALITY

1 Introduction

Natural kinds are a particular locus of debate concerning modality. Modern essentialists claim that natural kinds are the subjects of necessary but aposteriori truths and have real essences. Here I look at these claims and the founding arguments for them, as presented in the work of Kripke and Putnam. From where do these arguments ultimately derive their support? From intuition? Or do they have an independent foundation in semantic theory? And if intuition is the foundation, is this intuition reliable?

2 Kripke on natural kinds, necessity, and essence

In *Naming and Necessity* and elsewhere, Saul Kripke makes a number of claims:

(I) For some natural kinds, A and B, it is a necessary aposteriori truth that A is B.

(II) By investigating the basic explanatory properties of a kind, science attempts to find its nature and so essence.

(III) Natural kind terms are rigid designators, designating natural kinds.

(IV) Where 'A' and 'B' are natural kind terms, 'A is B' expresses an identity.

Note that claims (I) and (II) and are in the material mode—they concern truths about kinds. Whereas claims (III) and (IV) are in the formal mode—they are about the properties of linguistic items such as names.

Let 'A' and 'B' be natural kind terms. Then according to (III) they rigidly designate natural kinds. So 'A is B' must express an identity between natural kinds, the proposition that natural kind A is natural kind B (which Kripke calls a 'theoretical identification'). This is not trivial, for if 'A' and 'B' were non-rigid definite descriptions, to be analzyed as Russell proposes, then 'A is B' would not express an identity. So (IV) is a consequence of (III). Now consider the general claim that identities are necessarily true and are sometimes known aposteriori. This supports (I), if we think that what goes for identities in general goes for identities concerning natural kinds in particular. That is, we should accept (I) if we think that identities among natural kinds are not a special subset of identities that are always knowable apriori (if true).

Examples of (I) are as follows (page references are to Kripke 1980):

Gold is the element with the atomic number 79 [116]
Water is H_2O [116, 128]
Heat is the motion of molecules [99, 129]
Temperature is mean molecular kinetic energy [129]

Light is a stream of photons [116, 129]
Flashes of lightning are flashes of electricity [132]

Kripke also includes in his discussion some propositions that do not strictly conform to (I), such as:

Cats are animals [123, 125]
Whales are mammals [138]
Lightning is an electrical discharge [116]
Light is a form of electromagnetic radiation [129]

These look to be examples of a claim weaker than (I), *viz.*:

(V) For some natural kinds, A and B, it is a necessary aposteriori truth that all members of A are members of B.

We ought also take note of other claims that Kripke makes that are weaker than (I) and (II):

(VI) For some natural kinds, e.g. A, and properties, e.g. P, it is a necessary aposteriori truth that all members of the kind A have property P.

(VII) By investigating the basic explanatory properties of a kind, science attempts to find its essential properties.

Claim (VII) is weaker than (II) because (II) claims that we can find something that is *the* essence of a natural kind, implying that possession of the essence is sufficient as well as necessary for kind membership—a full or complete essence. Whereas (VII) identifies only essential properties which are necessary but may not be sufficient for kind membership—a partial essence. So, for example, the essence of gold might be the property of being comprised solely of atoms with 79 protons in the nucleus—which is both necessary and sufficient for something being gold. Whereas there might not be a property that is *the* essence of being a cat; even so, it is essential to being a cat that an entity is an animal.

3 Kripke's argument

Although Kripke starts his discussion of natural kinds by mentioning the theoretical identifications we find in (I), his arguments do not focus on those identifications directly. Rather he starts by arguing that:

(i.a) Certain properties we might use to characterise a natural kind and which might be thought to be components of the definition of the kind term, are not necessary properties of the kind.

These properties are ones that have often in the subsequent literature been called 'superficial' properties—these being properties that are readily observable and which are apparent to those with little or no scientific training. He also argues that:

(i.b) The superficial properties in the alleged definitions of the kind terms are not jointly sufficient for kind membership.

Kripke then argues that:

(ii) Some other (non-superficial) properties that we find aposteriori to characterise the kind are necessary to it.

These properties are typically ones discovered by sophisticated scientific means. They have often been called 'structural' properties, that is inner, structural properties of member of the kind. While Kripke does often talk in terms of structure, it is not clear that Kripke thinks that all such properties are structural. (And subsequent authors—see McGinn 1976—have extended Kripke's essentialism about origin to biological kinds.)

Regarding (i.a) Kripke considers a supposed definition of 'gold' as 'yellow metal' (which he finds in Kant). If correct, that would make being yellow a necessary condition on a piece of matter being gold. He then asks how we would respond if it were discovered that the yellow appearance of the stuff being dug out of gold mines is an illusion, that the stuff is in fact blue. If the definition of gold as yellow metal is correct, then it would have to be said that the stuff being dug in the mines or of which many rings are made and so on is not gold—that what we had been calling 'gold' is not gold after all. Kripke denies that this would be our response. Instead we would say that we have discovered that gold is not yellow after all, but is blue.

Kripke presents an analogous argument regarding the proposal that tigers are, by definition, large carnivorous quadrupedal felines, tawny yellow in colour with blackish transverse stripes and white belly. If so there could not be three-legged tigers. But there is no contradiction in the idea of a three-legged tiger. Furthermore, it could be that all those who had ever seen the relevant animals had been subject to an illusion making it appear that they have four legs whereas in fact they all have three legs. Such a discovery would not lead us to say that there are no tigers. Rather we would say that tigers are three-legged animals, not quadrupedal.

Kripke also considers whether the properties in the proposed definitions of 'gold' and 'tiger' are sufficient for membership of the kinds. We should pause to note that when Kripke talks of the proposed definitions he uses the indefinite article (as one naturally does): 'gold is *a* yellow metal' 'a tiger is *a* large carnivorous quadrupedal feline ...'. These could be read as implying that the properties listed are *not* to be taken as jointly sufficient for kind membership. If gold is *a* yellow metal, then perhaps there are other yellow metals that are not gold. Still, I think there is a reading on which the definiens properties are intended as jointly sufficient. Kripke assumes as much and then argues that something could satisfy the definiens without being of the kind. So something could have all the identifying features of gold without being gold—iron pyrites (fool's gold) he suggests. Iron pyrites is not a very good example, since the definition we are working with is 'yellow metal' and iron pyrites is not metal, nor does it have the readily observable properties of a metal (being ductile, malleable, etc.). Nonetheless we can understand the reference to fool's gold as intending some substance, real or imaginary, that has all the relevant allegedly defining 'superficial' properties of gold. Indeed, there is such as substance if the definition is 'yellow metal', since brass is a yellow metal; indeed a kind of brass, Muntz metal (famous for sheathing the hull of the Cutty Sark), is also known as 'yellow metal'. Clearly Muntz metal, an alloy of copper, zinc, and iron, is not gold. Likewise, if we discovered some hitherto unknown animals that resemble tigers in all the allegedly defining respects, but which were not mammals but reptiles, we would not say that some tigers are reptiles, but that these animals are not tigers.[1]

[1]For this purpose, one has to take 'feline' not to name a biological category but to mean 'cat-like in appearance'.

Arguments (i.a) and (i.b) establish that for some kinds at least, the 'superficial' properties by which we very frequently identify members of a kind are not properties that that are individually necessary or jointly sufficient for kind membership. So which properties are necessary and/or sufficient for kind membership? Kripke says that given that it is true that gold has atomic number 79, nothing that fails to have atomic number 79 would be gold. However similar something is to gold, if we found that it lacks this property we would say that it is not gold. So having atomic number 79 is necessary to being gold. It is clear also that scientists had to discover that gold has atomic number 79, so this is a necessary aposteriori truth about the kind. This establishes claim (VI).

Similar arguments Kripke takes to establish (V). Consider a non-actual world where there are cat-like entities that are in fact demons (1980: 126). It would not be that case that these are cats that are not animals. Rather they would not be cats at all. So, Kripke concludes, it is necessary that cats are animals.

Now consider again the case of gold and the property of having atomic number 79. The arguments so far considered show that being made entirely of stuff with atomic number 79 is necessary for being gold. That is consistent with there being other stuff that is not gold but which also has atomic number 79, that is, with there being multiple types of stuff with atomic number 79, only one of which is gold. While Kripke does not directly argue that having atomic number 79 is sufficient for gold, the many arguments for (i.a) support this conclusion. Consider any property distinct from having atomic number 79 that one might consider necessary for being gold (e.g. being yellow in colour). By considering stuff with atomic number 79 but without these properties (e.g. that the stuff we dig out of gold mines is a blue metal) we conclude that none of these is necessary but that having atomic number 79 is necessary. But if no other property is necessary for being gold, then it is reasonable to conclude that having atomic number 79 is sufficient for being gold. So having atomic number 79 is necessarily both necessary and sufficient for being gold. This establishes (I).

4 Essentialism

Kripke calls identities that are instances of (I), e.g. that gold is the element with atomic number 79, 'theoretical identities' because they were once unknown truths that scientists had to hypothesize and then confirm with appropriate experiments and observations. Propositions of the forms in (V) and (VI) are likewise theoretical claims established by aposteriori scientific investigation. That much is clearly right. For (VII) to be true it also needs to be the case that these aposteriori necessary properties, sub-kind relationships, and identities characterise features of the kind that are (a) explanatory in a basic way, and (b) its (partial) nature and essence. For (II) to be true, instead of (b) we need (b′) its (full) nature or essence.

It looks correct to say that the properties etc. in question are explanatory in a fundamental way. The atomic number of an element determines its electron structure, which in turn determines is chemical behaviour, its colour, electrical conductivity, and so on. So atomic number is explanatory in a basic way that these other properties are not. That lightning is a form of electricity explains why it can burn what it strikes, why it can disrupt electrical equipment, why it can be attracted to a metal lightning rod, why it will charge a Leyden jar, and so on. So (a) is true.

What about (b)? In his brief argument for (II) (and (VII) Kripke does talk of scientists discovering the 'very nature' of the substance (Kripke 1980: 124). But he does not at that point talk about the essence of the kind. Later he does so, saying that, 'In general, science attempts, by investigating basic structural traits, to find the nature, and thus the essence (in the philosophical sense) of the kind' (1980: 138). But he does not really argue for these claims about nature and essence. It is as if he thinks that once one is persuaded of the relevant aposteriori necessities, it will be obvious that they concern the nature and essence of a kind.

Very plausibly our intuitive notion of 'nature' supports that. What is the nature of gold? It is the element with atomic number 79. What is the nature of water? H_2O. What is the nature of lightning? A form of electric discharge. What about essence? In *Naming and Necessity* Kripke does not discuss essentialism in detail. Essential properties are necessary properties. But are all necessary properties essential properties? Kripke does not say. Even so, the quoted sentence suggests that in his view essential properties are those necessary properties that reflect the nature of the kind in question. So a plausible hypothesis to attribute to Kripke says that the concept of essence is much like the concept of nature, except that it also includes a modal commitment: whatever is an essential property of a kind X is necessarily necessary for being an X; whatever is the essence of a kind X is necessarily sufficient for being an X. This would fit with Fine's (1994) distinction between necessity and essence. Not all necessary properties are essential: it is a necessary but not essential property of *Socrates* that he is a member of the singleton set containing Socrates, whereas is an essential property of that *set* that Socrates is a member of it. For Fine the difference between necessity and essence is that essence concerns the nature or identity of the entity in question. And so we can establish the essentialist elements of (b) and (b′), since the properties etc. in question not only concern the nature of the kinds, they do so necessarily.

5 Putnam on natural kinds and modality

In 'The meaning of "meaning"', Putnam puts forward arguments that have conclusions very similar to Kripke's. Putnam (1975: 232) is far more explicit than Kripke that he holds these modal conclusions to be consequences of a semantic theory regarding natural kind terms:

> What Kripke was the first to observe is that this theory of the meaning (or 'use' or whatever) of the word 'water' (and other natural-kind terms as well) has startling consequences for the theory of necessary truth.

The theory of meaning and the related claims about modality are developed primarily in Putnam's famous 'Twin Earth' though experiment. We imagine that there is another planet in our galaxy, Twin Earth, that is with respect to its superficial properties identical to our planet Earth. Some of its inhabitants speak a language almost indistinguishable from English and they refer to the colourless, odourless, thirst-quenching liquid that fills the seas, lakes, and rivers of Twin Earth as 'water'. It turns out that this liquid, superficially just like water, has a complicated chemical formula quite unlike 'H_2O' which we may abbreviate by 'XYZ'. Putnam asserts:

> (a) On Twin Earth the word 'water' means XYZ.

whereas

(b) On Earth the word 'water' means H_2O.

This has consequences for what the inhabitants of the two planets say: on Earth some people (those with an education in chemistry) will say 'water is H_2O' whereas their Twin Earth counterparts will say 'water is XYZ'. However, if we turn the clock back to 1750, before the chemical revolution, such differences will disappear: the utterances of Earth-dwellers and their Twin Earth counterparts will be indistinguishable, and, says Putnam (1975: 224), they will be exact duplicates in 'appearance, feelings, thoughts, interior monologue, etc.'. The members of such pairs are 'in the same psychological state'.

Assuming that the extension of 'water' as used on Earth has not changed between 1750 and the present day (and so likewise on Twin Earth), this gives Putnam his principal conclusion: it is not the case *both* that for every term there is a meaning/intension that determines its extension, *and* that knowing this meaning/intension 'is just a matter of being in a certain psychological state'. This conclusion is semantic externalism. The argument for that conclusion depends on there being a reasonable sense of 'psychological state' such that the Earth-dwellers and their Twin Earth counterparts are in the same psychological state. Some might assert that any reasonable sense of 'psychological state' is such that these psychological states are themselves to be understood externalistically. Still, it will remain the case that two individuals can use words that have differing extensions despite being in identical brain states: 'Cut the pie any way you like, meaning just ain't in the head' (Putnam 1975: 227).

Putnam maintains that we would assent to (a) and (b) and that this is explained by his theory of the meaning of 'water' and other natural kind terms, which is as follows. He supposes that we give a meaning explanation for such terms by use of a particular sample of the kind in question—for example a glass of water—and of a cross-world 'same kind relation'.[2] Such a meaning explanation may proceed by pointing to the glass of water and declaring 'this is water' while intending the following (where w ranges over possible worlds, and x ranges over particulars):

> (S) $\forall w \, \forall x$ in w (x is water $\leftrightarrow$ x bears the same kind relation to the entity referred to by 'this' in the actual world).

It is intending this that makes 'water' a rigid designator.

Putnam assumes (but does not explicitly state) something like:

> (MC) For many substance kinds such as water, something bears the same kind relation to a sample of the substance if and only if it has the same microstructural constitution (e.g. chemical composition) as the sample.

If the stuff in the glass on Earth is H_2O then (S) and (MC) imply:

> (N) In all possible worlds, something is water if and only if it is H_2O.

And because (S) is part of a meaning explanation for 'water', we can say not only that water is H_2O, we can also say that 'water' means H_2O. Hence (b) is true (and, *mutatis mutandis*, so is (a)).

It is an aposteriori truth that water is H_2O. So (N) and its derivation imply something related to (I):

[2] Putnam talks of a 'same liquid relation', but that introduces avoidable difficulties concerning kinds, including water itself, whose samples may exist in gaseous or solid states as well as the liquid state.

(I)′ For some natural kinds, A and B, it is a necessary aposteriori truth that something belongs to A iff it belongs to B.

This is an example of what Mackie (2006) calls 'predicate essentialism'.

6 The semantics of natural kind terms and modality

Putnam's argument appears to derive the essentialist conclusion (I)′ from his theory of meaning, and the quotation at the beginning of section 5 implies that he thinks that Kripke makes a similar derivation. Salmon (1979) argues that no such derivation can be made—attempts to do so rely on hidden essentialist assumptions.

Consider the argument for (N). This depends crucially on (MC), which Putnam does not discuss but just takes as given. Without (MC) there is no argument for (N). Consider an alternative to (MC):

> (SC) For many substance kinds such as water, something bears the same kind relation to a sample of the substance if and only if it shares the same superficially observable characteristics (e.g. appearance, taste, aroma, etc.) as the sample.

If (SC) were true then we could not derive (N). Samples of XYZ would also be water. For that matter, there might even be possible worlds with slightly different laws of nature where H_2O does not have the superficial properties it has in the actual world. If so some possible samples of H_2O are not water. So (N) would be false as would be (I)′.

Where does this leave the argument for essentialism? Salmon (1979) thinks that Kripke's essentialist claims are independently plausible, whereas others, such as Mellor (1977), who take a similarly negative view of the derivability of essentialism from semantics, regard this as a reason to reject essentialism—what its supporters take to be the key argument for essentialism is unsound.

It is far from clear that Kripke—unlike Putnam and despite what Putnam says—does think that the argument for essentialism must proceed via a semantic route. As I have described it, Kripke's discussion starts with (i.a) and (i.b) before concluding with the essentialist (ii). Since (i.a) and (i.b) are the propositions that one would need (and which Kripke uses) to defeat a Fregean account of meaning, it is perhaps natural to suppose that what he is doing in *Naming and Necessity* is, first, arguing against Fregeanism about meaning, then, secondly, arguing in favour of an alternative semantics in which the idea of rigid designation plays a major role, and, thirdly, extracting from the latter an argument for essentialism. A different view of Kripke's strategy is that in establishing (i.a) and (i.b) he is providing premises from which, along with further plausible premises, the rejection of Fregeanism will follow. That in turn motivates the alternative semantics. But the argument for (ii) is independent of that alternative semantics. That proceeds naturally in the wake of the arguments for (i.a) and (i.b). For if the superficial properties of a natural kind do not characterize what that natural kind is, which, if any, properties do characterize it?

In favour of this interpretation is the fact that in the pages discussing (i.a), (i.b), and (ii) there is no in depth discussion of semantics, and none of his own alternative account of reference and meaning. Kripke does not present his essentialism about natural kinds as deriving from his semantic theory. Just as I have outlined it above, his argument is simple and straightforward. He asks 'Given that gold *does* have the

atomic number 79, could something be gold without having the atomic number 79?' He then considers counterfactual circumstances in which a substance, however similar to gold in superficial respects, was not an element with atomic number 79. And he declares that we would not say that such a substance is gold. Far from using semantic theory, Kripke appeals to our philosophical 'intuitions' about counterfactual situations in the manner that is familiar throughout philosophy (Williamson 2007).

So what of Putnam's argument, which Putnam does seem to think depends on semantic theory? That argument's principal aim is also to reject Fregeanism, so inevitably it draws on semantic theory. The essentialist conclusion is presented as a corollary. So the key question is, which parts of the overall argument does the essentialist corollary require? Note first that if essentialism was all that Putnam was interested in, he could have proceeded just like Kripke, asking 'Given that water *does* have the chemical composition H_2O, could something be water without being H_2O?'. The Twin Earth thought experiment provides us with the counterfactual example, in which we judge that the substance on Twin Earth, for all its superficial similarity to water, is not water because its chemical composition is XYZ, not H_2O. That argument, exactly parallel to Kripke's, does not use anything not found in Putnam's argument. It may not appear this way because the consequences of the thought experiment on which Putnam concentrates are (a) and (b), which are metalinguistic rather than object-language claims—these are supposed to the linguistic discoveries that the Earthian and Twin-Earthian travellers make. Putnam's discussion makes it clear that (a) and (b) are much the same as:

> (a′) On Twin Earth the extension of the word 'water' is XYZ (and so not H_2O).

whereas

> (b′) On Earth the extension of the word 'water' is H_2O (and so not XYZ).

However, (b′) could not be asserted if one thought that a sample of XYZ were water. I suggest then that although (b′) is a metalinguistic claim, it also embodies the judgment that the counterfactual XYZ would not be water, and it is this judgment that delivers the essentialist conclusion. Putnam is not explicit that this judgment is one that we would make. But he himself makes that judgment when he sets up the thought experiment in saying, 'I shall suppose that the oceans and lakes and seas of Twin Earth contain XYZ and not water, that it rains XYZ on Twin Earth and not water, etc.' (1975: 223). Putnam *should* have said 'I shall suppose that the oceans and lakes and seas of Twin Earth contain XYZ and not H_2O'. Assuming that he had said the latter, it would then be a matter of an additional non-linguistic judgment that the oceans etc. of Twin Earth do not contain water. From that judgment (along with trivial linguistic propositions) claim (a′) follows. Independently, the essentialist claim also follows, that necessarily XYZ is not water, and we are at least close to the essentialist target, that necessarily water is H_2O.

Nonetheless, Putnam does think that (a)/(a′) and (b)/(b′) are explained by his semantic theory. So cannot essentialism *also* be derived from the semantic theory? As we saw above these judgments are explained by three components of the semantic theory: (i) pointing to a sample that in fact is H_2O (on Earth; but XYZ on Twin Earth) and declaring 'this is water'; (ii) intending (S); (iii) the proposition (MC). But for the essentialist conclusion, we do not need (i) and (ii). (N) also follows from (MC) plus:

(S′) $\forall w \forall x$ in w (x is water $\leftrightarrow$ x bears the same kind relation to canonical water-samples in actual world.

and the fact that canonical water samples in the actual world are H_2O. These propositions are all in the object language. So again, it is nothing specific to the semantic theory that is doing the work in generating essentialism. Clearly it is (MC) that is doing the heavy lifting. As mentioned, Putnam does not explicitly assert (MC) but it is clearly implied by what he says.

The conclusion we should draw is that essentialism cannot be derived from the core of the semantic theory. But that need not undermine the case for essentialism. For Kripke's argument (as discussed in section 3) does not use the semantic theory at all, depending instead on intuitions concerning the kinds in question. And although Putnam's argument looks as if it uses the semantic theory, in fact it does not, depending instead on a general claim about kinds that must seem to him so obvious and unremarkable that it does not need stating explicitly at all. The defender of essentialism should therefore say that the argument for essentialism is simple—it just is intuitively highly plausible, as revealed either by counterfactual thinking or by our willingness to make tacit essentialist assumptions.

At the outset I noted that (I) is supported by (III), and that too looks like a claim about essentialism derived from a semantic claim. The response is the same: an argument for essentialism from rigid designation cannot provide independent grounds for asserting essentialism. For the very considerations that support the view that 'water', 'gold', and 'tiger' are rigid designators are those that directly support essentialism without a detour through semantics. It is because our counterfactual thinking does not track the superficial characteristics of these kinds that essentialism is plausible. And it is because our counterfactual thinking does not the track superficial characteristics of these kinds that it is plausible that 'water', 'gold', and 'tiger' each designate one kind across all possible worlds, where sameness of kind is a matter of the sameness of basic explanatory (usually structural) characteristics.

Although Salmon is right that essentialism cannot be derived from the semantics of natural kind terms, the discussion in the preceding paragraph shows that the semantics and modal metaphysics of natural kinds are not entirely independent. The point can be extended to suggest that rather than essentialism being derivable from semantics, it is semantics that is sensitive to modal metaphysics. Let K be a natural kind of the sort covered by (MC). So whether a particular belongs to the kind K depends on whether it has the right microstructural constitution. But the microstructures of things and kinds are properties that we can be entirely ignorant of and may not even be able to conceptualise—in 1750 people were not able to have the concept 'possessing the microstructure H_2O'. The Fregean theory requires that the extension of any term is determined by properties that the user of the term is able to conceptualise. So Fregeanism implies that we can only have a term that refers to K if we know what the microstructure of K is. So, in particular, (some) speakers today can use 'water' to refer to the kind whose microstructure is H_2O—because we can use 'having microstructure H_2O' as an intension that will determine the correct extension. But no speakers in 1750 could do that.[3] More generally, if essentialism is true about natural kinds, Fregeanism implies that scientifically ignorant speakers cannot refer

[3]Conceivably they could do so if they used an intension which is necessarily co-extensive with 'having microstructure H_2O'. It is an interesting question whether could have been such an intension available to speakers of English before the chemical revolution.

to many natural kinds.[4] Consequently, if one thinks that essentialism is true and that the scientifically ignorant can refer to these natural kinds, then one must reject Fregeanism.

7 Objections to natural kind essentialism

One line of attack on Kripke–Putnam natural kind essentialism takes as its target the semantics of natural kinds terms that Kripke and Putnam develop—Mellor (1977) is a prime example of this approach (although Mellor has other arguments besides). Since, as I have argued, the support for natural kind essentialism comes from elsewhere than the semantic theory, I shall not outline such objections here. The three principal further kinds of response are:

> (i) To reject the allegedly simplistic theoretical identifications upon which Kripke and Putnam rely. In particular, science tells us that it is not strictly correct to say that water is H_2O.
> (ii) To argue that we do not in fact classify samples/items in a manner consistent with essentialism.
> (iii) To question the reliability of the intuitions on which the arguments for essentialism are based.

We should not be essentialists, because water is not H_2O

Kripke and Putnam argue that the essential properties of many natural kinds are properties detailing some aspect of the microstructure of examples of the kind. The theoretical identities mentioned in (I) are often an identity formed by using a name of a kind on the one hand and a microstructural description on the other hand: water is H_2O, gold is the element with atomic number 79, and so on. A number of philosophers of science have argued that the science on which these identities is based is too simplistic. The microstructure of water just is not H_2O. Needham (2011) and van Brakel (1986), for example, argue for this on several grounds. First 'H_2O' is not a microdescription at all, since what it tells us is that water is "a compound of hydrogen and oxygen in the proportions 2 moles to 1", which is a macroscopic description (Needham 2011: 9). Secondly, the microstructure of liquid water is very different from that of water vapour and that of ice. So there is no such thing as *the* microstructure of water. Furthermore, they argue, certain clearly macroscopic descriptions of water *do* suffice to identify water, for example thermodynamic properties such as the triple point (the temperature and pressure at which the gaseous, liquid, and solid phases of a substance co-exist). Thirdly, not all H_2O is water, because D_2O, deuterium oxide, is a species of H_2O but we would not and should not regard it as water (LaPorte 2004). This is because D_2O has different rates of reaction from the rates of reaction of H_2O and this has significant consequences—D_2O can be toxic for some animals.

While these observations may require essentialists to be clearer about how their view should be articulated with respect to the particular case of water, they do not

[4] This in turn implies that if, being scientifically informed, we do now use terms such as 'water' to refer to natural kinds, then there has been meaning change over time. In that sense although essentialism is true, knowledge that what we refer to as 'water' has the essence 'H_2O' is apriori knowledge resulting from a semantic decision to change the meaning of 'water'. Something like this is Joseph LaPorte's (2004) view.

seem to undermine the case for microessentialism in general. Nor do they unequivocally refute microessentialism about water in particular (Hendry 2006). Because it is chemically complex water presents particular problems in getting the details right that other chemical substances do not. Indeed Needham (2011: 4) accepts that gold's atomic number 79 is a sufficient microstructural condition for identifying gold. First, to be told that water is a compound of hydrogen and oxygen is to be told something about its microstructure, that it contains atoms of oxygen and hydrogen and not any of kind of atom. Secondly, although the different phases of water do have different microstructures, that is not to say that there is nothing that those microstructures have in common: for any sample of pure water, whether solid, liquid, or gas, is composed (almost) entirely of H_2O molecules.[5] Thirdly, there are good reasons for thinking that D_2O is a species of water (after all it is called heavy water). While reaction rates are different for D_2O from 'normal' water, this is the same for any compound where there are isotopic variants for the constituent elements (which is to say: almost every compound).

We are not essentialists when we classify things as water

A distinct line of criticism of essentialism argues that what we regard as water is not determined by its microstructure. Some items that are largely H_2O are regarded as water, such as sea water and tap water, but others that are no less composed of H_2O are not regarded as water, such as tea and Sprite (Malt 1994; Chomsky 1995). Abbott (1997) responds first by remarking that we do say that tea is composed mostly of water precisely on the basis that it is mostly composed of H_2O. She proposes that our different classifications of impure substances is just another example of vagueness, whereby we tolerate a degree of imprecision that is appropriate to the context. And it does seem correct that we do regard impure substances as water if their differences from pure water are irrelevant or negligible in the context—mineral water and herbal tea may be similar in their proportion of non-water constituents, but in the context of request for water with a meal in a restaurant, the impurities in the mineral water that make it differ from pure water are irrelevant but corresponding differences between tea and pure water are relevant in that context.

We are essentialists, but this is just a cognitive bias

Philosophers' use of intuition as evidence for or against philosophical theories of all kinds has been subject to criticism (e.g. Swain et al. 2008). The intuitions to which Kripke and Putnam appeal have been equated with *psychological essentialism* (Leslie 2013). If psychological essentialism is an unreliable source of intuition, then we should not readily rely on Kripke–Putnam type intuitions, despite a strong human tendency to do so. Psychological essentialism is the thesis that humans have a cognitive disposition to believe that the underlying intrinsic structures and mechanisms of things are better predictors of behaviour than more superficially observable characteristics and it is such underlying mechanisms that determine the nature and identity of a thing. Numerous studies show that this cognitive disposition is found in young children. Whether such findings are sufficient to undermine

[5] The 'almost' here refers to the fact that in liquid water a small proportion—fewer than 6 per billion at 37°C—of the molecules dissociate into OH^- and H_3O^+ ions. Hendry (2006) regards this dissociation as the biggest obstacle to simple microessentialism. Hendry regards water as what you get when you bring H_2O molecules together—what you get includes the ions produced by the dissociation caused by bringing the molecules together in a liquid.

Kripke–Putnam intuitions depends on whether those cognitive dispositions are reliable. Leslie points to the rejection of microessentialism about water (as discussed above) as well as similar arguments against microessentialism about species as evidence for the unreliability of these cognitive dispositions. Note that the intuitions that Kripke and Putnam appeal to are not simply 'water has an essence that is its internal structure'. Rather they are, strictly, intuitions of the form, 'given that scientists tell us that all and only water around here is H_2O, then something would not be water if it is XYZ'. Now that intuition is not shown to be wrong by the finding (if correct) that it is false that scientists tell us that all and only water around here is H_2O. More generally, it has not been shown that the cognitive dispositions underlying such intuitions are unreliable when supplemented by philosophical reflection and correct scientific information.

8 Conclusion

I have outlined some of the founding arguments for essentialism about natural kinds above. Like so many philosophical arguments, these depend on 'intuitions'—quasi-intuitive judgments about counterfactual cases. Although it looks as if Putnam and (to a much lesser extent) Kripke think that the essentialist conclusions can be derived from semantic premises, this is an illusion since those semantic premises themselves contain substantive essentialist commitments. One could take this, as I do, to show the deep-seated nature of those intuitions. One could instead argue that this is wrong—we do not have such deep-seated intuitions; indeed our commonsense classificatory practices are incompatible with essentialism. Even if one accepts that such intuitions are there, those does not show that they are reliable. They may be just the product of cognitive biases or of scientific ignorance (or of both working together).

References

Abbott, B. 1997. A note on the nature of "water". *Mind* **106**: 311–19.

Chomsky, N. 1995. Language and nature. *Mind* **104**: 1–61.

Fine, K. 1994. Essence and modality. In J. Tomberlin (Ed.), *Philosophical Perspectives 8: Logic and Language*, pp. 1–16. Atascadero, CA: Ridgeview.

Hendry, R. 2006. Elements, compounds and other chemical kinds. *Philosophy of Science* **73**: 864–75.

Kripke, S. 1980. *Naming and Necessity*. Oxford: Blackwell.

LaPorte, J. 2004. *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press.

Leslie, S.-J. 2013. Essence and natural kinds: When science meets preschooler intuition. In T. S. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology*, Volume 4, pp. 108–65. Oxford University Press.

Mackie, P. 2006. *How Things Might Have Been*. Oxford: Oxford University Press.

Malt, B. C. 1994. Water is not H_2O. *Cognitive Psychology* **27**: 41–70.

McGinn, C. 1976. On the necessity of origin. *Journal of Philosophy* **73**: 127–35. This paper is principally about the essentiality of the origin of biological individuals, following Kripke's *Naming and Necessity*. He briefly also mentions species as having their origins essentially. This presages views in the philosophy of biology that biological taxa may have non-intrinsic, historical essences (see Okasha 2002 in *Natural Kinds in Biology*). It also reminds us that Kripke's essentialism doesn't require microstructuralism.

Mellor, D. H. 1977. Natural kinds. *British Journal for the Philosophy of Science* **28**: 299–312.

Needham, P. 2011. Microessentialism: What is the argument? *Noûs* **45**: 1–21. This paper lays out the problems for the microessentialism of Kripke and Putnam.

Putnam, H. 1975. The meaning of 'meaning'. In *Mind, Language and Reality: Philosophical Papers, Volume 2*, pp. 215–71. Cambridge: Cambridge University Press.

Salmon, N. U. 1979. How not to derive essentialism from the theory of reference. *Journal of Philosophy* **76**: 703–25.

Swain, S., J. Alexander, and J. Weinberg 2008. The instability of philosophical intuitions: Running hot and cold on truetemp. *Philosophy and Phenomenological Research* **76**: 138–55.

van Brakel, J. 1986. The chemistry of substances and the philosophy of mass terms. *Synthese* **69**: 291–324.

Williamson, T. 2007. *The Philosophy of Philosophy*. London: Wiley-Blackwell.