

Kuhn on Reference and Essence

Alexander Bird
University of Bristol

Résumé : La thèse kuhnienne de l'incommensurabilité semble mettre en cause le réalisme scientifique. Une réponse à cette mise en cause consiste à se focaliser sur la continuité de la référence. La théorie causale de la référence, en particulier, semble offrir la possibilité d'une continuité de la référence susceptible de fournir une base pour l'espèce de comparabilité entre théories que requiert le réaliste. Dans « baptiser et rebaptiser : la vulnérabilité des désignations rigides », Kuhn attaque la théorie causale et l'essentialisme auquel cette théorie est liée. La position de Kuhn est défendue par Rupert Read et Wes Sharrok [Read & Sharrok 2002b]. Dans cet article, j'examine les arguments présentés par Kuhn, Read et Sharrok, et je montre qu'ils ne fournissent aucune raison de douter ni de la théorie causale, ni de l'essentialisme.

Abstract: Kuhn's incommensurability thesis seems to challenge scientific realism. One response to that challenge is to focus on the continuity of reference. The causal theory of reference in particular seems to offer the possibility of continuity of reference that would provide a basis for the sort of comparability between theories that the realist requires. In "Dubbing and Redubbing: The Vulnerability of Rigid Designation" Kuhn attacks the causal theory and the essentialism to which it is related. Kuhn's view is defended by Rupert Read and Wes Sharrock [Read & Sharrok 2002b]. In this paper I examine the arguments presented by Kuhn, Read, and Sharrock and show that they provide no reason to doubt either the causal theory or essentialism.

1. The challenge of incommensurability and the referentialist response

Kuhn's incommensurability thesis tells us that there is an important difficulty in translating between the scientific terminology of theories from either side of a scientific revolution. Even where the terminology looks the same — the same words are employed — their meanings have changed, so that they no longer express the same concept. Considered as two homophonic terms, the one term does not translate the other.

Prima facie this is unwelcome to realists. One of the ways in which science progresses is by correcting the mistakes of earlier scientists. We once thought the Earth to be at the centre of the universe, orbited by the planets, whereas we now know that the Earth is not at the universe's centre, but is itself a planet orbiting the Sun; it was once thought that water is an element, but we have since discovered that it is a compound; the idea that new species have evolved has displaced the old mistaken view that species are immutable. And so on. But if the new theories and those they replace do not mean the same things by the terms they use, it looks as if we cannot straightforwardly say that the later theory denies what the earlier one asserts, in which case we cannot say that it represents a correction to and improvement on the earlier theory.

However, on further reflection the realist does have a way of defending the idea that later science genuinely corrects (and so rejects) some parts of earlier science. The realist can say that whatever else may have changed, the *reference* of the key scientific terms has remained constant. We use our scientific terms to talk about things, kinds, properties, quantities and so forth. This 'talking about', a relationship between words or phrases and things in the world is the relationship of reference. And reference is what is important to truth (and falsity). For it is the properties of the stuff water that determine whether or not it is an element. Continuity of reference therefore allows later scientists to say things that are incompatible with what earlier scientists had said, where 'incompatible with' means 'not true at the same time as'. Boerhaave used the term 'aqua' in his *Elementa Chymiae* (1732), Lavoisier used 'eau' in his *Traité élémentaire de chimie* (1789), and Berzelius used 'vatten' in his *Lärobok i Kemien* (1808-18 to 1843-48). They ascribed different and incompatible properties to what they were talking about. Various accounts of reference makes it both possible and plausible that nonetheless they were referring to the same substance and hence that the later chemists were able to correct their predecessors.

The idea that reference is a key concept in assessing the progress of science and that there has been a large measure of referential continuity over time, even over revolutions, I shall call *referentialism*. Although it is not their view, Read and Sharrock nicely encapsulate the importance of referentialism to realism thus:

There is only one world, and the Realist (unlike the Relativist or Idealist) can have coherent things to say about our deepening knowledge of that (one) world, because we (all of us, especially via scientists) are always in touch with the world. We are in touch with it through our naming of bits of it, and our growing knowledge about the nature of what we name. It may that at different times people had radically wrong ideas about the nature of the world — they may have thought, for example, that water was a primitive element — but, through being in contact with bits of the world (e.g. with water), and through naming it, they always had some kind of ‘basis’ to their claims. We can connect with them and what they said, because there are direct (albeit long) chains of connection — linguistic and (more generally) ‘causal’ — between their use of these words and our use of them. They may have ‘meant’ something very different by their words, but the reference of their words was just the same as our reference for the same words. Reference — and in particular a causal theory thereof — will settle the problem which meaning poses. [Read & Sharrok 2002, 153].

If referentialism is correct, the challenge presented by Kuhn’s incommensurability thesis is rather weaker than at first appeared. As Read and Sharrock continue, the realist thinks that ‘Kuhn is wrong, because he thinks that incommensurability of meaning is important and deep. It is, in fact, completely shallow and unimportant, once one understands that the real reference for natural kind terms remains continuous over time and through ‘revolutions’.’

Given such a consequence for Kuhn’s thesis, it was natural that Kuhn should reject referentialism. Kuhn thus argues against it in his paper “Dubbing and Redubbing: The Vulnerability of Rigid Designation” [Kuhn 1990a] and his view is defended and elaborated by Rupert Read and Wes Sharrock [Read & Sharrok 2002].

This paper aims to defend the referentialist view expounded above so very clearly by Read and Sharrock against the criticisms that they and Kuhn level against it. They also criticize the related thesis of *essentialism*, and I shall also defend this.

The aim of this paper is thus not to criticize Kuhn’s incommensurability thesis. Rather it is to show how the referentialist and essentialist can mount an effective defence against the specific criticisms of Kuhn,

Read, and Sharrock. If such a defence is successful, then the thesis of incommensurability may be thought to present no interesting challenge to the realist. The realist can regard incommensurability and realism as consistent. Of course, there are other reasons one might have for doubting realism, and these may even be part of a more broadly Kuhnian framework. These issues will not be my concern here. My concern is limited to articulating a defence of the referentialist versions of realism against the specific criticisms raised by Kuhn, Read and Sharrock. Mounting this defence does not require therefore that I defend realism in general. This will be important because the referentialist defence of scientific realism against the apparent threat of incommensurability is itself thoroughly realist. For example, the referentialist defence takes it that there is a world of independently existing entities studied by science and that we are able to refer to such entities. The referentialist defence is easiest to express if we assume that scientists can get to know things about such entities, although I do not believe that this assumption is an inevitable part of the referentialist defence. That referentialism makes realist assumptions is no failing, since the context in question is the defence of realism, not the larger issue of whether realism is superior to anti-realism.

2 Referentialisms and essentialism

Read and Sharrock regard the position they describe as a ‘kind of essentialism’. This is misleading, for although essentialism is closely bound up with questions of reference, it is an entirely distinct view from referentialism. A property which is essential to a kind is a property such that no entity could be a member of that kind without possessing that property¹. So ‘being mammalian’ is an essential property of the kind *horse* since nothing could be a horse without being a mammal. Historically essences have been held to come in two varieties, nominal and real essences. A *nominal* essence of a kind *K* is a property that an object belonging to the kind possesses in virtue of the verbal definition of the term ‘K’. The *real* essence of a kind derives instead from the nature of the kind. The presupposition of the idea of a real essence, whereby a kind can have an essential nature, is that the kind in question has a real, natural existence. One could regard some random collection as the members of some kind defined for that purpose (e.g. the kind *K* is defined as having exactly, the Eiffel tower, my desk, and the number nine as its

¹This may only be a necessary condition of being an essence, not a sufficient one. See [Fine 1994].

only members, the kind L is defined as comprising all and only objects that are either cubes or are coloured red). Such kinds could be regarded as having nominal essences (necessarily something is L iff and only if it possesses the property of being either red or cubic). But clearly such kinds do not have a nature distinct from their nominal essences, and so they do not have real essences distinct from their nominal essences.

In its most general form, *essentialism* is the view that there is a non-trivial distinction between essential and non-essential (accidental) properties. However, a more specific conception of essentialism takes it that there is a non-trivial distinction between *real* essences on the one hand and on the other hand firstly accidental properties and secondly merely nominal essences. Thus on the more specific conception there exist natural kinds with real essences and not in every case are these real essences identical to or immediately derivable from the nominal essences and the verbal definitions of the kind terms². This kind of essentialism may be extended to incorporate real essences of natural properties and individuals also. Since the features constituting the real essence derive from the nature of a kind (property, individual) rather than from its verbal definition, and because the nature of a kind may be hidden from us, revealed only by *a posteriori*, scientific investigation, it may be the case, hold many essentialists, that in some instances a real essence may be uncovered only by empirical investigation and cannot be known *a priori*.

In what follows, I shall reserve the term ‘essentialism’ for the more specific thesis, that there is a non-trivial conception of real essences.

Since referentialism is a thesis about words while essentialism is a thesis about things and their metaphysics, it should be clear that they are not the same thesis. Indeed referentialism as I have characterised it is consistent with the denial of essentialism. A term ‘K’ may continue to refer to a kind K over time without its being the case that all Ks share a real essence. This is true of the artificial kinds K and L introduced above. Members of those kinds share a nominal essence but not a real essence. Referentialism *per se* does not require the existence of *natural* kinds.

One reason for wanting to divorce referentialism from essentialism is that within referentialism there are different views about how reference works and while some of these are closely associated with (although not

²An essentialist of this kind might maintain that the real essence reflects a kind’s *real* definition. But then the real definition and the verbal definition need not be the same.

identical to) essentialism, other views about the working of reference are not so associated. I shall discuss the two main versions of referentialism: (a) the Fregean version, employing a concept of sense which determines reference, and (b) the Kripke-Putnam version, employing a causal or non-Fregean, externalist conception of reference determination. Only the latter is closely related to essentialism.

3 Fregeanism and the first referentialist response

One way of expanding on referentialism as characterised above is to employ Frege's famous sense-reference distinction [Frege 1892]. According to this view a term has associated with it a *sense* (*Sinn* or *intension*). What exactly Fregean senses are or ought to be is a matter of dispute, but the following aspects of sense are central:

- (i) the sense of a term determines its reference (*Bedeutung*, *extension*);
- (ii) the senses is grasped by someone who understands the term in question.

A common way of thinking about sense is as a set of descriptions such that (i)* something is the reference of the term if and only if it satisfies those descriptions, and (ii)* someone who understands the term believes those descriptions to hold of the reference of the term. Of course a speaker may have many beliefs about the reference of a term she uses, and not all those beliefs will be part of the sense of the term. The beliefs associated with the sense will be true *analytically* and will be knowable *a priori*. For example someone might propose that the sense of 'human' is 'rational animal'. If that is correct then since sense determines reference, all and only rational animals will be humans, and because senses are grasped by speakers, everyone who understands the term 'human' will believe (perhaps only tacitly) that humans are rational animals; the latter proposition will be true analytically and will be knowable *a priori*.

One feature of the Fregean view that will be important later is a consequence of the combination of (i) and (ii). Let T_1 and T_2 be terms with different references. Then, by (i), T_1 and T_2 must have differing senses. By (ii) we may expect the beliefs of a subject with respect to the reference of T_1 to be different from their beliefs with respect to T_2 . (Strictly, the range of analytically true, *a priori* known, beliefs of the subject will differ). Conversely, two subjects who are identical in all respects, differing at most in ways about which they are entirely ignorant

(e.g. differing in aspects of their environment beyond their ken) will use terms with the same sense and so the same reference.

The Fregean view provides a referentialist response on behalf of the realist to Kuhn's incommensurabilist challenge. Several terms may have different senses but share the same reference. In Frege's example, 'Venus', 'Hesperus', 'Phosphorus', 'the morning star', 'the evening star' are all terms that refer to one and the same planet; but, arguably, they have different senses. This means that the realist can accept Kuhn's claim that the meanings of terms change through revolutions so that a word used after a revolution does not translate the same word as used before the revolution. For translation requires preservation of sense. So the realist will say that there is translation failure, because sense is not preserved through a revolution. Nonetheless, because terms with different senses can have the same reference, the reference may well have remained the same. And it is plausible to think that typically reference is indeed preserved — or at least that scientists intend reference to be preserved — because scientists very often want to add to beliefs and correct beliefs *about the same things that they were previously interested in*.

Two things should be noted. First, the Fregean view shows that incommensurability and referentialism (and so realism) are consistent, only so long as incommensurability is a thesis about change in senses/intensions. Were it the thesis that reference changes, then Fregeanism would be of no help. Secondly, even if incommensurability and referentialism are consistent, that fact does not prove that there is indeed continuity of reference. I shall make some remarks about each of these in turn.

The precise nature of Kuhn's incommensurability thesis is a matter of debate, not least because it changed over time. However, it is clear that it is a thesis primarily about meaning in a sense that is much closer to Frege's *Sinn* than his *Bedeutung*. For example, incommensurability is often characterised as the thesis that terms from distinct paradigms (e.g. from either side of a scientific revolution) may fail to translate one another. Translation requires preservation of sense (*Sinn*, *intension*). So incommensurability as untranslatability is a thesis about a change in the sense of a term. The incommensurability thesis, at least in its earlier incarnation, is a consequence of the widespread theoretical-context view of the meaning of theoretical terms. The latter is also a thesis concerning the sense of those terms, not their reference. There are two further, related reasons for thinking that the incommensurability thesis concerns sense rather than reference. First, a corresponding thesis concerning Fregean reference (*Bedeutung*) is much less plausible than a thesis concerning sense. *Prima facie*, it is plausible that when there is a scientific

revolution the shift in theories, practices and so on means that words acquire a new sense, that they cannot directly translate their predecessors, and that there are barriers to perfect communication. But the idea that a later theory is referring to different kinds and entities from its predecessor is much less plausible. Often, it is true, a new theory might initiate reference to an additional set of kinds and entities. So the discovery of the structure of DNA permitted talk of new entities (base pairs, for example). But substances being referred to beforehand were still being referred to afterwards (DNA itself, for example). In other cases, old terms that failed to refer are given up (e.g. phlogiston), but in such cases also, this is against a background of continued reference (to air, water, iron, iron calx, muriatic acid, and so forth). Indeed if there were significant discontinuity in reference one would wonder why the phenomenon should be thought of as a scientific revolution as opposed to simply a change in subject matter (such as a change from physics to zoology). Since the idea of incommensurability as a shift in reference is at least *prima facie* implausible, it would need powerful arguments to support it. Kuhn does not provide such arguments — what he does say provides much better support for the incommensurability thesis understood as a change in something akin to sense.

The second reason Kuhn would not have thought of incommensurability as a shift in reference is that reference (as Fregean *Bedeutung*) is the sort of realist notion that Kuhn either eschewed or rejected. Kuhn rejected the ideas of truth and increasing nearness to the truth on the basis that they employed an illegitimate notion of what is ‘really there’ [Kuhn 1970a]. But the function of reference is precisely to make a connection between our words (and thoughts) and what is really there. Now it is true that Kuhn does say the following, when discussing the common proposal that “Apparently, Newtonian dynamics has been derived from Einsteinian, subject to a few limiting conditions.”:

[T]he physical referents of these Einsteinian concepts [spatial position, mass, time] are by no means identical with those of the Newtonian concepts that bear the same name. (Newtonian mass is conserved; Einsteinian is convertible with energy. Only at low relative velocities may the two be measured in the same way, and even then they must not be conceived to be the same) [Kuhn 1970a, 101-102].

This passage is puzzling if Kuhn’s ‘physical referent’ is taken to be the same as Frege’s *Bedeutung*. For then the passage would imply that there exists in the world both of two distinct quantities, Newtonian mass and Einsteinian mass — and that Newton as talking about one and Einstein

about the other. But that is implausible — why should we think that the world contains both of these two different quantities? And if there really is Newtonian mass in the world, doesn't that mean that Newton was right after all (about 'Newtonian mass')? No, either there is no mass as referred to by Newton (the term 'mass' as used by Newton is like 'phlogiston' as used by Priestley) or there is mass as referred to by Newton — it is the same as the quantity referred to by Einstein, only that Newton had some false beliefs about it (the term 'mass' as used by Newton is like 'Mars' as used by Ptolemy). This passage becomes much less puzzling if we do not take 'physical referent' as *Bedeutung*, but rather to mean something like 'what scientists took themselves to be referring to' or 'what scientists believed about the intended reference'. This would give us an 'internal' sense of reference, so that in this sense a term can have a physical referent, even if there is no physical quantity in the world corresponding to it. Furthermore, the 'physical referent' of a term, in this sense, will change if there is a theoretical shift (for the simple reason that what scientists believe about the intended reference changes). It might be that Kuhn thinks of incommensurability as a change in 'reference' understood in this sense. But then the incommensurability thesis does not, on its own, represent any threat to scientific realism (which is the issue under discussion), for the realist can accept that there is a change in 'physical referent' understood in Kuhn's internalist way (just as the realist can accept a change in sense) while continuing to assert that there can be continuity in reference to a genuine quantity or entity in the world.

It is important to be clear about what the referentialist claims the Fregean position would show, if true. It shows that it is possible for sense to change (leading to untranslatability) while reference is preserved. It does not *guarantee* that reference is indeed preserved. There could be cases where the sense associated with a term T at some time determines reference R_1 but after a revolution the sense associated with T has changed and now picks out a different entity, R_2 . Thus the Fregean version of referentialism is consistent with the view that there is discontinuity of reference. Therefore, all that the Fregean view would show, if correct, is that continuity of reference and so scientific realism are consistent with incommensurability (interpreted as a shift in something like sense). The appeal to Fregean referentialism is thus only a *defence* against an incommensurabilist critique of realism; it does not on its own show that the realist view must be correct.

Could the incommensurabilist critique be supplemented to show that revolutionary shifts in intension are also accompanied by shift in refer-

ence? I have already suggested that any argument to the effect that there are shifts in reference would have to be more powerful and detailed than anything Kuhn has supplied. I shall return to the issue of reference shifting later at the end of §4. More plausible on a Kuhnian version of Fregean referentialism is that there will be reference failure rather than discontinuity of successful reference.

Consider the possibility that the sense of T at time t_1 involves very many beliefs — $\{B_1\}$. As a result of a revolution the sense of T changes significantly, so that at a later time t_2 , the beliefs — $\{B_2\}$ — constituting the sense of T are very different. Since the sets $\{B_1\}$ and $\{B_2\}$ are both large and significantly different, there is a considerable danger that the same things will not satisfy both sets, and hence that reference is not preserved. Even so, this will not involve a genuine shift in successful reference. If the earlier set $\{B_1\}$ is large, constituting most or all of the scientists' beliefs about T at t_1 , then there it is likely that *nothing* will satisfy $\{B_1\}$ and so T has no reference and T will not exist. And indeed we may expect this to be the case, since we are interested in scientific revolutions. Given that there is a revolution involving a change in beliefs, precipitated by anomalies or other kind of failure of the scientists' beliefs at t_1 , that suggests that indeed the beliefs $\{B_1\}$ cannot all be true, in which case T cannot refer.

This outcome, whereby there is reference failure (as opposed to shift in successful reference) is the result of combining a Fregean view with the contention that the sense of a term is very 'thick', involving a large number of beliefs. What is significant is that during the period preceding the writing of *The Structure of Scientific Revolutions*, the standard view about the meaning of scientific terms was that all parts of a theory in which a term is employed contribute to its meaning. So all the beliefs encapsulated in a theory would contribute to the term's sense, making it very thick indeed.

In so far as we are still concerned with genuine, external reference, such a position is unattractive. It is unattractive to the realist, since widespread reference failure is in tension with the initial characterisation of realism as being committed to the improvement in beliefs about the same things. (At the same time it is also incompatible with the view that there are shifts in genuine external reference. But this position need not be unattractive to the 'internal reference' view discussed above). So for the realist, the Fregean version of referentialism ought optimally be supplemented with arguments as to why the sense of the relevant

scientific terms is likely to be thin and not thick³.

It is worth pointing out that the Fregean view had troubles of its own. One source of trouble was Quine's influential argument that the analytic-synthetic distinction is empty. If Quine were right the whole notion of Fregean sense is misconceived, for one could not distinguish between those propositions or beliefs that constitute the sense of a term and those that do not. A rather different source of trouble for the Fregean view came from Kripke and Putnam [Putnam 1975a] [Kripke 1980]. Read and Sharrock suggest that they developed their ideas 'in part specifically in opposition to this idea of Kuhn's'. (The idea of Kuhn's in question was that the world changes as a result of a scientific revolution, which can be understood as the claim that scientific revolutions generate taxonomic incommensurability). Although Putnam and Kripke did not have world-changes or incommensurability as their primary target when they formulated their views, it is true that they and realist philosophers of science who were interested in such matters realised the significance of this version of referentialism.

4 Kripke, Putnam, and the second referentialist response

What Kripke and Putnam were primarily concerned to combat was the combination of the central Fregean claims, (i) and (ii) above. They held that it is not the case that reference is determined solely by 'what is in the head' of a speaker, such as the speaker's beliefs. Hence nothing can satisfy both (i) and (ii), and so reference is not fixed by a Fregean sense. This is established by Putnam's famous twin-earth argument. The key idea is that it is conceivable that two subjects O_1 and O_2 can be internally exactly alike yet when O_1 uses a term T it has a reference that is different from the reference of the same word T used by O_2 . What does differ between O_1 and O_2 is their environment — but the differences in their environments are differences of which they are unaware. Thus it would appear that a set of beliefs or a Fregean sense is not sufficient to fix reference. Something else, something external to O_1 and O_2 , something in their environment of which they are ignorant, plays a part also. In Putnam's argument we consider two worlds E and TE . E is the actual world while TE is a world exactly like E , except that where E has H_2O TE has a distinct compound XYZ that nonetheless shares the same

³For a detailed discussion of the relationship between Fregean intensionalism and incommensurability, see [Bird 2000, 163-178].

superficial qualities as H_2O (smell, taste, boiling and freezing points, ability to support life, etc.). Consider the year 1950. O_1 uses the term ‘water’. In so doing O_1 refers to H_2O but not to XYZ (it is a truth about the actual world that water is H_2O , and not XYZ). Since E and TE are symmetrical we must say that O_2 in TE when using the term ‘water’ refers to XYZ and not H_2O . *Ex hypothesi* E and TE are alike except for the difference between H_2O and XYZ. And so whatever O_1 believes concerning what he calls ‘water’ O_2 also believes, concerning what he calls ‘water’. So their beliefs about ‘water’ do not distinguish them — they both believe that the following sentences are true: ‘water is important for life’, ‘rain and snow are constituted from water’, ‘sodium chloride dissolves in water’ and so forth, and there is no sentence such that one believes it and the other does not. If Fregeanism were true, then the sense of ‘water’ used by O_1 and the sense of ‘water’ used by O_2 must be the same — by (ii) in §3 above. And so, by (i) the references must be the same. But we have seen that the references are not the same. So Fregeanism is false.

The question is then raised, what is it that fixes reference if it is not a Fregean sense, if it is not a set of beliefs in the head of the speaker? A common answer, but not the only possible answer, is that it is some kind of causal connection between the speaker’s use of the term and the reference. Why is it that when using the term ‘water’ O_1 refers to H_2O but not to XYZ, while O_2 refers to XYZ but not to H_2O ? The suggested answer is that O_1 is causally connected to H_2O but not to XYZ and vice-versa for O_2 . O_1 acquired the use of the term ‘water’ when as a child his parents said ‘this stuff is water’ pointing to the stuff coming out of the tap when O_1 was having a bath. That stuff, being in the world E , was H_2O . In the corresponding learning situation, O_2 ’s parents pointed to stuff that came out of the bath-tap on TE , which is XYZ.

This allows a referentialist response to incommensurabilism that is different from the Fregean one. The referentialist can say that whatever else changes during a scientific revolution, the causal connections between the scientists (and their use of words) and the things in the world do not change. Their beliefs may change radically, but what those beliefs are about will not change, because that relation, the referential relation, is not fixed by beliefs but by a causal connection that remains in place. O_1 may grow up to have all sorts of strange beliefs about water and these may change a great deal, but they are still all beliefs about water, the stuff that is H_2O , because the causal connection between his use of the term ‘water’ and the stuff water (H_2O) was established before

and independently of those beliefs⁴.

Because on the causal view of reference-fixing, reference is fixed independently of beliefs, the causal view does not fall prey to the objections to the Fregean view. For example, it doesn't matter much if one wants to say that much or all of a theory is part of the 'meaning' of some scientific term, because any such conception of meaning cannot be one where meaning determines reference. And it is reference that is significant in being able to say that a later theory corrects an error in an earlier theory, or in saying that later scientists know more about some subject matter than did earlier scientists. As regards the realism debate, the idea of meaning becomes insignificant, as does the incommensurabilist thesis that there is meaning change of a kind that precludes translation⁵.

It should be noted that a pure causal theory of reference (only the causal connection counts — theoretical beliefs play no part), is not the only account of reference consistent with the Kripke-Putnam rejection of Fregeanism. Mixed accounts will do also, and there are other accounts where a causal connection is not always required.

It is intriguing to note that at one point Kuhn himself took a very similar view:

The distinction between a theoretical and a basic vocabulary will not do in its present form because many theoretical terms can be shown to attach to nature in the same way, whatever it may be, as basic terms. But I am in addition concerned to inquire how "direct attachment" may work, whether of a theoretical or basic vocabulary. In the process I attack the often implicit assumption that anyone who knows how to use a basic terms correctly has access, conscious or unconscious, to a set of criteria which define that term or provide necessary and sufficient conditions governing its application [Kuhn 1974, 467 footnote 11].

Here Kuhn seems to be rejecting a Fregean approach to reference and endorsing one that allows for a direct attachment to nature (i.e. not via a set of criteria or descriptions), of the kind that Kripke and Putnam promote. However, as we shall see, Kuhn later rejected and attacked the Kripke-Putnam approach. It is interesting to ask why Kuhn's view changed. I speculate that Kuhn was not conscious of a change and that

⁴What would be required for a change in the reference of O_1 's term 'water' would not be a change in O_1 beliefs, but a change in O_1 's causal connections with kinds of stuff. So if O_1 were moved, unknowingly, to TE or a TE -like place, his new causal contact with XYZ may mean that in due course O_1 might be referring to XYZ with the term 'water' and no longer to H_2O .

⁵For a discussion of the causal theory and its variants in relation to incommensurabilism see [Sankey 1994a] and [Bird 2000, 179-181].

the difference in view represents a difference in intended contrast, a difference in date of writing, and a difference in emphasis of research. The passage quoted is a footnote to a 1974 paper on paradigms, which were the focus of his early work. One of the key elements in the paradigm idea is to oppose the view that scientific deliberation is a matter of employing rules. In extending this idea to the role of paradigms in conferring meaning on terms, it is natural to claim that meaning is not specified by rules (e.g. correspondence rules, criteria, necessary and sufficient conditions etc.) but instead may be conferred other means, for example, by direct attachment. Later, however, Kuhn's focus turned much more towards incommensurability, and in this context the idea of direct attachment looks unwelcome since it would undermine incommensurability in the way described, and hence is appropriate for attack in the 1990 paper [Kuhn 1990a].

Earlier I mentioned that the Fregean account of referentialism does not guarantee continuity of reference. It is possible for the sense connected with a word to change sufficiently that its reference changes also. Nonetheless, the dialectical relevance of Fregean referentialism is that it shows how incommensurability and scientific realism *can* be consistent. Thus, as it stands, incommensurability does not threaten scientific realism, if the Fregean view is right. To show that it does, one would additionally need to show that the relevant shifts in sense were sufficient to yield shifts in reference also (but without involving reference failure). Similar remarks may be made with regard to the version of referentialism based on the causal theory of reference. There can be continuity of causal connection, and so continuity of reference, even though there are changes in the theory associated with a term. So again, while some version of the thesis of incommensurability (e.g. concerning the impossibility of translation) might be true, it would still be consistent with scientific realism⁶. That is enough to remove the immediate threat posed by incommensurability. The causal theory of reference does not itself guarantee that there will always be continuity of reference. Even on this theory it is possible for the reference of a term to change. It has been said that the name 'Madagascar' was originally applied to the mainland of Africa by those who lived there, but European explorers misunderstood their interlocutors and applied it to the large island that now bears that name. The use of the name in connection with the island by those explorers and their successors forged a new causal connection between the name

⁶It might be mentioned, however, that the causal theory of reference, in making do without the notion of sense, casts some doubt upon there being a problem of translatability.

and the island and it is this use we have inherited. Now, it is conceivable that something like this could happen in science. One could argue that Democritus' 'atom' has a different extension from ours⁷. But plausible cases are few and far between. What is required is that there is an established causal link to one reference or extension and that a new causal link is forged to another reference or extension. Although Kuhn mentions 'redubbing' in the title of his paper, he nowhere argues that such shifts actually take place. To make out a case just the like the 'Madagascar' case, one would have to show that the later scientists mistook the intentions of the earlier scientists. That is a possible scenario, but on the whole one might judge it to be unlikely. Elsewhere he does assert that there have been changes in extension. For example, he argues that the extension of 'planet' has changed, because before Copernicus the Sun was held to be a planet but the Earth was not, whereas now we take the opposite to be true. But this of course does not establish any shift of extension rather than a shift in what the extension was thought to be [Bird 2000, 160-162]. We are often wrong about extensions, and we often disagree about them.

To conclude, the causal theory of reference does not itself guarantee continuity of reference. It is conceivable that revolutionary changes in science result in shifts of reference and extension of the 'Madagascar' variety. But it would take a lot more historical detail than Kuhn has offered to make out such a case. As it stand, the causal theory shows how there *can* be continuity of reference despite even revolutionary scientific changes, and so shows how scientific realism is not threatened by the thesis of incommensurability as Kuhn formulated it.

5 Criticising Putnam

The causal theory of reference and Putnam's argument that gives rise to it have come in for considerable discussion and criticism. I shall consider here only criticisms related to our discussion of Kuhn.

The critic might first remark that it is unlikely that O_1 and O_2 have analogous beliefs. After all (on E) it is widely known that water is H_2O (and correspondingly it will be known on TE that water is XYZ). If O_1 and O_2 share in this widespread knowledge then they will have distinct beliefs, whereas the argument requires that they have isomorphic beliefs. The most straightforward reply is simply to stipulate that in the example

⁷Although one might doubt where the extension of Democritus' use of the term is established by a causal connection, rather than via a sense.

O_1 and O_2 do not have this knowledge. Although they may know that substances are elements or compounds made up of elements in chemical combination, they may not have had enough education to know which compound water is. The objector might respond by suggesting that what, for O_1 , fixes H_2O as the referent of water, is the belief, widespread in O_1 's society on E , that water is H_2O . This would still abandon the letter of Fregeanism, since it would be no state of O_1 that fixes the reference of 'water'. Nonetheless, it might be said to retain the spirit of Fregeanism, since the proposal is, in effect, that senses exist but are social rather than individual. A standard way of cutting through all of this discussion is to turn our attention to the ancestors A_1 and A_2 of O_1 and O_2 respectively, who inhabit E and TE in 1750. Since this date is before the discovery of the chemical constitution of water, the beliefs of A_1 and A_2 are isomorphic as are the beliefs of A_1 's society and A_2 's society. There is no reason to suppose that the reference of 'water' has changed between A_1 and O_1 — what we refer to as water is the same stuff as what our ancestors in the eighteenth century referred to as water. Similarly A_2 and O_2 are referring to the same stuff. Thus A_1 is referring to H_2O and A_2 is referring to XYZ. And so once again we have a difference in reference without any difference in belief, and so without a difference in sense — whether individual or social.

Kuhn's primary response to Putnam is to point out that the existence of XYZ is chemically impossible — at least, it is incompatible with what modern chemistry tells us. Let us define 'water-like' thus. Something is 'water-like' if it has the superficially observable properties of water (it is a tasteless, colourless substances, boils at $100^\circ C$, freezes at $0^\circ C$, support life etc.) Thus according to Putnam's story, XYZ is water-like but is not water. Kuhn points out that unless modern chemistry is badly wrong, it is not possible for there to be a water-like substance that is not H_2O . There can be no water-like XYZ, unless XYZ is H_2O after all. So the envisaged scenario is impossible, as far as we know. But were it possible, then modern chemistry is wrong, and hence the discovery on TE that water is XYZ would prompt a scientific revolution. Consequently, even if O_2 were untouched by this, the (scientific) society which surrounds him would be very different from that surrounding O_1 .

I find it difficult to see what the relevance of Kuhn's response is supposed to be. The causal theory and essentialism have both been placed under close and critical scrutiny by analytic philosophers. But Kuhn's arguments have been ignored even by the most stringent critics of Kripke and Putnam. In my view this is because Kuhn seemed to have a tin ear for the sort of philosophical argument that is standard in modern phi-

losophy (though less common in philosophy of science than elsewhere). Nonetheless, Read and Sharrock have sought to resurrect Kuhn's arguments. They say "[Kuhn] is raising a fundamental difficulty for the essentialist/referentialist view. He is saying: 'If we try taking Putnam's example seriously it turns out that Putnam does not really offer up a tenable thought-experiment after all.'" [Read & Sharrock 2002b, 154].

Let us reflect on what Putnam's example is supposed to be and do. Read and Sharrock call it a 'thought-experiment'. If it is a thought-experiment, it is a thought-experiment not in chemistry but in the philosophy of language. It is intended to show us how particular words and expressions function. In particular it is designed to show what does and what does not fix the reference of our natural kind terms. It does this by getting us to reflect on a particular imaginary scenario, and eliciting our judgments about what the kind terms used by the imaginary people in that imaginary scenario refer to. The 'data' yielded by the thought-experiment are those judgments (that O_1 refers to H_2O while O_2 refers to XYZ). Those judgments license the inference that our implicit understanding of the function of kind terms is that they do not have their references fixed by a speaker's beliefs. Since we are competent users of the English language we may be expected to have a reliable implicit understanding of the function of various categories of expression in English and so we may infer that in fact kind terms do not have their references fixed by a speaker's beliefs.

Does Putnam's imaginary scenario have to be a possible scenario for it to do its work? No it does not. Putnam is not telling us that XYZ is possible. He is asking us to imagine that it is possible. Remember that the relevant data are our judgments about what the imaginary O_1 and O_2 are referring to in the imaginary scenario. So for the thought-experiment to do its work it is necessary only that the scenario be imaginable. And here 'imaginable' need be taken only in a very weak sense. We want to elicit judgments from ourselves and others in response to the imaginary scenario, and for that we need only *think* that the scenario is possible. Even if in fact it is impossible, that is no obstacle to our imagining that it is possible or to entertaining the scenario. And there are certainly conceptions of 'imagine' that allow this. For it is perfectly coherent to imagine even mathematical impossibilities being possible. For one can, in the relevant sense, imagine its being possible to trisect any angle. Say one were doing a study concerning what mathematical propositions people find interesting or surprising. It is perfectly coherent to try to elicit a judgment from someone by asking, "Imagine that you could trisect any angle. Would you find that an interesting or surprising possibility?"

It is no objection to such a study that it is impossible to trisect the angle. Nor is it at all relevant that if it were possible to trisect the angle, then modern mathematics would be badly wrong and discovering that one could trisect the angle would lead to a revolution in mathematics.

6 Defending Putnam

Although the above is an entirely sufficient response to Kuhn's comment, I also suggested that in order to bypass all such considerations we could imagine Putnam's scenario being discussed in the early to mid nineteenth century [Bird 2000, 183]. The idea is this. The discovery that water is H_2O was a development that emerged in the first half of the nineteenth century, with the most important steps being taken in the period between 1804 and 1820, thanks initially to Dalton and then to Berzelius. However, at that time chemistry was in a sufficiently nascent state that it is plausible to suppose that Berzelius and his contemporaries could not be sure that there is no distinct compound that shares the superficial properties of water. We now have to imagine some forerunner of Putnam carrying out his thought experiment in 1850. We elicit the same judgments, that in using the term 'water' O_1 on E is referring to H_2O while O_2 on TE is referring to XYZ , even though the beliefs they have are the same (remember that O_1 and O_2 do not know much chemistry). However this time, it is not open to a putative forerunner of Kuhn's to object that there could be no such XYZ that has the superficial properties of water but is a different compound.

Read and Sharrock respond to this proposal as follows:

But this just seems incoherent: it seems that now there no longer is a thought-experiment. For how are we supposed to know that water is H_2O if we cannot rule out that some water is XYZ , if we cannot rule out XYZ as a starter in the game? And we cannot rule out the latter unless (e.g.) we know that XYZ cannot have the surface properties of water (...). Unless and until one has a good scientific reason for insisting water's extension cannot include XYZ , one will not need to endorse Putnam's 'externalist' conclusion. Kuhn's point is: such good scientific reason is 'only' in fact given *by the taxonomic etc. structure of modern chemistry*⁸.

This argument is not entirely clear. But it seems to be saying the following:

⁸See [Read & Sharrock 2002, 154] (emphasis in original).

- (1) to know that water is H_2O one must know that no XYZ can be water-like (i.e. no substance other than H_2O can have the surface properties of water);

and hence:

- (2) on my assumption that in 1850 they did not know that there is no water-like XYZ, they did not know then that water is H_2O ;

and so:

- (3) because in 1850 they did not know that water is H_2O they could not then have run Putnam's thought experiment.

If this is a correct interpretation, then the Read and Sharrock argument errs in three important respects: (i) it fails to understand Putnam's argument; (ii) it seems to assume precisely what Putnam's argument is intended to disprove; (iii) it contradicts what chemists and historians of chemistry themselves say.

(i) Read and Sharrock seem to think that by moving the thought experiment to 1850 I have undermined Putnam's thought-experiment. Whereas the experiment is possible in 1950 it is not possible in 1850. This is because in 1950 we knew there can be no water-like XYZ, which, *ex hypothesi*, we could not know in 1850.

But this is wrong. It plays no part in Putnam's thought experiment that we know that there can be no water-like XYZ. On the contrary, Kuhn's criticism is in part that Putnam *failed* to notice that according to modern chemistry there can be no water-like XYZ. The thought experiment works just as well in 1850, because those who run the experiment will be able to judge (or so Putnam and I expect) that individuals whose contact is with XYZ and not with H_2O will be referring by 'water' to XYZ and not to H_2O .

(ii) Read and Sharrock assume in (1) above that knowing that water is H_2O requires knowing that nothing other than H_2O has the surface properties of water. Why do they think this? The only reason I can think of is that they take the surface properties of water to be definitive of what water is. Something is water if and only if it has the surface properties of water. From this it would indeed follow that water is H_2O only if nothing other than H_2O can have the surface properties of water. But the idea that something is water if and only if it has the surface properties of water is precisely the idea that Kripke and Putnam are trying to

undermine. Remember that O_1 and A_1 know only about the surface properties of water (and correspondingly for O_2 and A_2 , concerning what they call 'water'). So if the reference of the term 'water' were fixed by their beliefs about water, then it would be fixed by beliefs about surface properties alone. But since, as the thought experiment aims to show, reference is not fixed by their beliefs, it follows that what water is cannot be defined solely in terms of its surface properties. So when Read and Sharrock assume in their criticism that water can be defined in terms of its surface properties alone, they assume what Putnam sets out to show is false. They beg the question against him.

(iii) Let us look more closely yet at the assumption (1) that knowing that water is H_2O requires knowing that nothing other than H_2O has the surface properties of water. That means that either (a) Berzelius (who died in 1848) did not know that water is H_2O , or (b) Berzelius did know that nothing other than H_2O could have the surface properties of water. But neither of these propositions is true. Berzelius certainly thought he knew that water is H_2O , and certainly historians of chemistry attribute that knowledge to him. But there is no reason to suppose that Berzelius ever gave serious thought to the question, could something other than H_2O have the surface properties of water, and there is even less reason to suppose that he did or could have known such a thing. In the quoted passage, Read and Sharrock seem to be saying that to know that water is H_2O , one must know that nothing other than H_2O is water-like, and that this knowledge must be based on good scientific reason, of the kind that only modern chemistry provides. (I presume from the context that 'modern' means twentieth or twenty-first century). But again that is false. As I have said, coming to know that water is H_2O , as Berzelius and his contemporaries did, did not require thinking about what other than H_2O could be water-like.

How did chemists come to know that water is H_2O ? The important work in showing that water is a compound of hydrogen and oxygen was carried out by Lavoisier, although similar experiments were carried out at the same time by Cavendish, Priestley, and Monge. The experiments were of two kinds: those that synthesized water, and those that analysed water by decomposition. In 1781 all the above-mentioned chemists found that the burning of 'inflammable air' (hydrogen) produced what they recognised as water. None seemed to ask whether water might be compounded of quite distinct substances. (Of course, water could be produced by other processes, such as the reaction of inflammable air with metallic calxes. But that just confirmed Lavoisier's view that the calxes contained oxygen). However, they could have had no theoretic-

cal reason for supposing that no other compound could produce a substance indistinguishable from water. (One hundred years later one could argue the point thus: to be water-like a substance must either be composed of small molecules, with strong inter-molecular forces, or of large molecules. All small molecules are known, and the only water-like one is H_2O . Larger ‘molecules’ are either crystals (including metals) or are organic compounds, all of which are not water-like. But such thoughts were not available to Lavoisier and his contemporaries). Quite probably there was an unstated background assumption that distinct compounds would have distinct observable properties, which implies that if the oxide of hydrogen is water-like, then nothing else would be. But such an assumption could hardly have been sufficiently well grounded to provide the ‘scientific good reason’ that Read requires for supposing that there is no other compound like water. Instead of relying exclusively on such an assumption, what clinches the identification of water as the oxide of hydrogen, is the dissociation of water by passing steam over heated iron. Lavoisier showed that if one passed steam through a red-hot gun barrel (and, in later experiments, over iron filings), the products were inflammable air and a calx of iron, just like iron ore.

The analysis of water by decomposition removed any room for doubt that the substance produced by burning inflammable air is water (although the chemists seem to have shown little doubt on that score). At the same time, the decomposition experiments were not by themselves conclusive for all (for example, Priestley hypothesized that the inflammable air had been contained in the iron). Of course, the synthesis of water from hydrogen and oxygen and the decomposition of water into hydrogen and oxygen, are experimental results that are logically compatible with the view that the extension of ‘water’ includes everything that has water’s surface properties. Thus those results are compatible with a view which says that there could be kinds of water that are not compounds of hydrogen and oxygen. It is instructive that neither Lavoisier nor any of his contemporaries considered such a possibility. The fact that *their* samples of water were a compound of hydrogen and oxygen licensed the conclusion they drew: water is a compound of hydrogen and oxygen. What is behind this license is not that they had good, specific, scientific reasons for supposing that there is in fact no other substance that has water’s surface features — for as we have seen, there were no such reasons. Rather, those chemists all took it that their samples as well as similar sources (i.e. the wells, rivers, rain etc. from which they might have obtained, by distillation, other samples of water) were samples of a single substance. *That* substance is water. The fact that they had not

ruled out the possibility of a water-like substance on Mars that is not a compound of hydrogen and oxygen was irrelevant, since that substance was not a potential source of a sample. They would have had no interaction with that substance. *A fortiori*, the fact that they had not ruled the possibility of synthesizing some water-like substance that might or might not exist in nature was irrelevant too, since that substance too was not a potential source of samples for them. It is true that Lavoisier et al. might have been mistaken in supposing all their samples and potential sources of samples to all be samples of the same substance. But notice that the risk of their being mistaken in this assumption is much less than the risk of being mistaken in the assumption that there is no possible substance that is water-like other than their samples.

To sum up: we may make two contrasting hypotheses about what Lavoisier et al. meant by ‘water’ (‘eau’ or ‘vatten’ etc.).

- (i) The extension of ‘water’ is precisely whatever is water-like. (In which case, in concluding that water is a compound of hydrogen and oxygen, they were making the highly risky assumption that there exists no other possible compound that is water-like and that no such other compound could be synthesized).
- (ii) ‘Water’ refers to that single substance of which all actual water-samples and potential water-samples (i.e. rain, wells, rivers etc. from which Lavoisier et al. could have obtained samples) are instances. (In which case, in concluding that water is a compound of hydrogen and oxygen, they were making the much less risky assumption that their water-samples and potential water-samples were all instances of a single substance).

The second of these hypotheses makes a much better explanation of why Lavoisier, Dalton and so on, reasoned as they did. First, we know that they did make the assumption ascribed to them in (ii). Indeed it is an assumption common to all those interested in the chemistry of water going back to Aristotle and before. But there is no evidence that they made the assumption ascribed in (i). Secondly, if we take it that the chemists in question were rational men, it is more plausible to attribute less risky assumptions in their reasoning than more risky ones.

Putnam’s thought experiment is one way in which we can see that the function of kind terms is to name a single kind, one with which we are interacting. But it is not the only way. The way in which people reason in actual circumstances is another source of argument. This is because people’s reasoning indicates their commitments and assumptions.

And since use of a term involves certain assumptions and commitments (e.g. concerning the extension of that term), reasoning using that term can reveal which assumptions an individual is making, and hence can reveal the function of a term. The reasoning of late eighteenth and early nineteenth century chemists reveals a commitment to a Kripke-Putnam view of natural kinds rather than a Kuhnian or Fregean one.

7 Essentialism

Earlier I separated referentialism from essentialism. The referentialism of Putnam is however related to essentialism. Fregean referentialism held that a term would pick out its referent by virtue of that referent satisfying the sense of the term. Putnamian referentialism holds that certain terms, including natural kind terms *name* their referents. This is an important difference. For consider the propositions: ‘George Orwell wrote political novels’ and ‘George Orwell is Eric Blair’. Clearly the former might not have been true had circumstances turned out differently. But the latter cannot but be true, however things might have been. Since ‘George Orwell’ and ‘Eric Blair’ are names of one and the same man, the sentence ‘George Orwell is Eric Blair’ simply says of a certain man that he is identical to himself. Since nothing can fail to be itself, under any possible circumstances, it is necessarily the case the George Orwell is Eric Blair.

The Putnamian view does not deny that the Fregean way of referring might be true for some terms — it just denies that it is true for all terms (specifically, it isn’t true for natural kind terms and proper names). Something like the Fregean view might be right for expressions such as ‘my favourite colour’, ‘the nearest planet to Earth’ etc.⁹ Such terms might have picked out different objects had matters been otherwise, and indeed they may pick out different objects at different times. The difference between terms that pick out the same objects under all possible circumstances and those that may pick out different objects, is marked by saying that the former are ‘rigid designators’. It is not only names of people and natural kinds that are rigid designators. Let us assume for illustration that a person could not have had different parents from the parents he or she does in fact have. Therefore, since ‘Auberon Waugh’ designates the same person under all possible circumstances, so does ‘the father of Auberon Waugh’, and the latter is a rigid designator also. Now consider ‘The father of Auberon Waugh is Evelyn Waugh’. Since this is true the terms ‘the father of Auberon Waugh’ and ‘Evelyn

⁹Russell, however, went on to deny that properly speaking these are referring expressions at all. Instead they are disguised forms of quantification.

Waugh' pick out the same man. Since those terms are rigid designators, they each pick out the same man in all possible circumstances. And so 'the father of Auberon Waugh' and 'Evelyn Waugh' pick out the same man in all possible circumstances. Hence 'The father of Auberon Waugh is Evelyn Waugh' is necessarily true. Looking at this another way, just as in the case of 'George Orwell is Eric Blair', the proposition says of a particular man that he is identical to himself, which of course cannot be anything but true.

The lesson of all this for natural kinds is as follows. We have already seen that 'water' functions like a name and so is a rigid designator. The same may be said, for the same reasons, of 'hydrogen' and 'oxygen' and of the chemical symbols 'H' and 'O'. The chemical expression 'H₂O' is clearly not a name in the ordinary sense. But like 'the father of Auberon Waugh' 'H₂O' is a rigid designator, since there is just one substance that can be a compound of hydrogen and oxygen in the ratio 2:1. Since it is true that water is H₂O, then 'water' and 'H₂O' designate the same substance, and since those terms are rigid designators, they designate the same substance under all circumstances. Hence 'water is H₂O' is a necessary truth. Note that in reaching this conclusion all that was assumed was (i) that water is H₂O and (ii) that 'water' and 'H₂O' are rigid designators, being names or functioning like names.

That it is necessary that water is H₂O lends support to the view that water has an essence, namely that it is H₂O. Locke tells us that the essence of something is some property such that something could not be that thing without possessing that property. And because necessarily water is H₂O, it is indeed the case that something could not be water without being H₂O.

The same conclusion can be reached in a slightly different way from Putnam's arguments. Those show that natural kind terms such as 'water' function as names, picking out a natural kind. So the extension of the predicate 'is water' includes all and only samples of that kind. What makes something a sample of a particular kind? There is room for some debate on this, but one thing is sure, and that is that members of the same kind must bear some deeper similarity than mere superficial resemblance. A robot cat would not be of the same kind as animal cats, however convincingly it behaved as a cat, since a creature with copper and silicon on the inside constructed by engineers does not belong to the same kind as an animal of flesh and bones begat in the standard way by its parents.

The clearest cases of entities all belonging to the same kind come from physics, since all electrons are *exactly* alike and all protons are

exactly alike and so on. So the kinds 'electron' and 'proton' are perfectly homogeneous. Moving on to chemistry we find that the kind terms have inhomogeneous extensions. This is because atoms and molecules might be in different states of excitation, and because of the existence of isotopes. Thus not all H₂O molecules are identical. Nonetheless they all belong to the same chemical kind. It is plausible to suppose that 'sameness of kind' may be fixed contextually. Thus what is meant by 'same kind' by chemists is fixed by their interests as chemists. Their interests do not extend to the differences between isotopes since these make no qualitative difference to chemical reactions. Thus to answer the question posed by Kuhn, 'is heavy water really water?' is 'yes it is' [Kuhn 1990a, 312]¹⁰.

There are other ways in which samples of what we ordinarily call water are not all the same. Much of what we call water contains considerable quantities of what is not H₂O. Seawater, dishwater, rain, mineral water and so forth all contain differing amounts of impurities. For this reason I stated the essentialist conclusion thus: 'in all possible worlds water consists (largely) of H₂O'. Read and Sharrock pick up on my use of 'largely' here. They ask whether essentialism would look so attractive if the impurities included XYZ, in significant proportions (say 25%) [Read & Sharrock 2002, 156]. They are right to raise this point. The primary response must be to refer to my proposal that the sameness of kind relation is contextual. In the chemical context, the impurities found in seawater, rain etc. would not be permitted and to the extent that something contains anything other than H₂O it is not a perfect sample of water. So if we are thinking of the *chemical* term 'water' then I was mistaken in including the parenthetical 'largely'. Chemical water is in all possible worlds *only* H₂O. On the other hand the extension of our everyday term 'water' does include these impure samples. We should conclude that the term 'water' as used by the chemist is a different term from the same word 'water' as used in everyday English. Interestingly, the extension of everyday 'water' does not cover samples that may have just as high a proportion of H₂O (tea, a tomato etc.). So we cannot say that everyday water is just impure chemical water. It seems that the extension of the everyday term is governed by a variety of practical considerations that may not be characterised straightforwardly. (Nonetheless, all these observations are consistent with the hypothesis that the practical considerations have the effect of modifying a core con-

¹⁰Kuhn also asks the corresponding question, 'is deuterium hydrogen?'. Again the answer is 'yes'. Physicists distinguish protium, deuterium, and tritium as the three isotopes of hydrogen.

ception of water, that it is a single chemical substance. Those complex, practical considerations would allow in samples with varying degrees of impurity but not all. If so, it would remain the case that necessarily everyday water is largely H_2O . That said, since we are concerned with the chemists' water, this is not strictly relevant).

What if our samples of water had always had a large admixture of XYZ? This would be significant, since, *ex hypothesi*, the XYZ would contribute to the dominant superficial character of all the samples of water, in a way that impurities do not. If our samples contained small amounts of XYZ, then that too might be considered as a case of an impurity. But a significant quantity of XYZ could not be so dismissed. In such a case, if it is discovered that water samples contained not only H_2O but also XYZ, then we would have to admit that pure water is not a single chemical kind but is a mixture of two compounds. And it would certainly be plausible to argue that water has no real essence (although some might reasonably argue that it is essentially a mixture of H_2O and XYZ). But the fact that the referent of 'water' might not have a real essence in the hypothetical case where our samples are a mixture of H_2O and XYZ does not show that in the actual case, where our samples are not such a mixture, that water is not essentially H_2O .

Cases such as the hypothetical mixture of H_2O and XYZ have been discussed before. The classic case is that of jade, which comes in two unrelated forms, jadeite and nephrite which are nonetheless superficially similar (though not entirely similar). In this case it is clear that there is no geological kind 'jade'; what we call jade is a mixture of kinds. But that does not mean that diamond, which does not have samples that are of different constitutions, is not essentially a form of carbon. The point can be seen most easily with respect to persons and the proper names referring to them (where Kuhn is willing to admit a causal theory and its consequences). Consider the following scenario: what one has taken to be an individual named John is not that but is a pair of identical twins who frequently substitute for one another. It may not be entirely clear what the name 'John' does: does it refer to some disjunctive entity or to the mereological sum of the individuals, or does it not refer at all and are sentences using the name 'John' without a truth-value? But what is clear is that the possibility of such a case does not impugn the causal theory of reference and the necessity of 'Eric Blair is George Orwell'. Similarly the case of jade and the hypothetical case of a mixture of H_2O and XYZ does not give us any reason to doubt the application of the causal theory to natural kinds and the necessity of 'all water consists of H_2O '.

8 Non-liquid water?

Kuhn also objects to Putnam's identification of water and H_2O on the ground that H_2O includes ice and steam whereas in 1750 'water' picked out only liquid water. Kuhn continues thus: "In 1750 the primary differences between the species recognized by chemists were still more or less those between what are now called the states of aggregation. Water, in particular, was an elementary body of which liquidity was an essential property." [Kuhn 1990, 311]. He goes on to say that the Chemical Revolution in the 1780s is what gave us a taxonomic transformation that allows a chemical species to exist in all three states. Read and Sharrock amplify Kuhn's comments, remarking, "It is only given our post-Lavoisieric framework that we are forced to see water as largely H_2O . Absent that framework, 'water'-in-all-its-states is not necessarily a natural kind." [Read & Sharrock 2002, 156].

These comments have very little impact on the essentialist argument for three reasons. Firstly, they are historically inaccurate. Phase transitions (changes of state) were widely regarded as changes in one and the same species long before Lavoisier. Secondly, the fact some chemists regarded water, ice, and steam as different chemical species does not show that the extension of their term 'water' has failed to include ice and steam. Thirdly, even if 'water' as used in 1750 did not include steam, essentialist conclusions may still be drawn.

Firstly, did scientists universally regard steam and water and different species before Lavoisier? The matter is far from as clear as Kuhn, Read, and Sharrock suggest. Aristotle tells us: "Now the sun, moving as it does, sets up processes of change and becoming and decay, and by its agency the finest and sweetest water is every day carried up and is dissolved into vapor and rises to the upper region, where it is condensed again by the cold and so returns to the earth." He also tells us that ice, snow, hail, are hoar-frost are water solidified by cold, that water freezes in winter, and that ice is made up of water¹¹. Locke starts by appearing to confirm Kuhn's view "If I should ask anyone whether ice and water were two distinct species of things, I doubt not but I should be answered in the affirmative; and it cannot be denied but that he that says they are two distinct species is in the right." [Locke 1690, III, vi, sect. 13]. But then he goes on to argue that someone brought up in Jamaica who found that in England a bowl of water froze over might regard the ice as hardened water and that "it would not be to him a new species, no more than congealed jelly, when it is cold, is a distinct

¹¹See [Aristotle *Meteorologica*, 354b27-30, 388b14, 347b36, 385b5].

species from the same jelly fluid and warm; or than liquid gold in the furnace is a distinct species from hard gold in the hands of a workman.” Locke draws a quasi-Fregean lesson, that the extensions of our species terms are fixed by nominal essences. Remember that for Locke the real essences that would distinguish between whatever real species there are, are undiscoverable. He need not have drawn such a conclusion; he could have held that whether water and ice really are the same or distinct species depends on scientific facts unavailable to him and to his contemporaries¹². What is important is that Locke does not regard it as absurd or demonstrably false to think that ice and water are the same species, and furthermore that he regards it as given that phase transitions for other substances (e.g. gold and jelly) are species-preserving, and so for Locke (and presumably his audience) it cannot be, *pace* Kuhn, that in general “the primary difference between the species recognized by chemists were still more or less those between what are now called the states of aggregation.” It is worth noting that the demonstration by Cavendish, Priestley and Lavoisier that water is a compound, by exploding hydrogen and oxygen together seems to *assume* that water-vapour and water are the same substance. For the product of combustion is water-vapour which must be condensed to give liquid water which was then subjected to confirming tests to show that it is water. If condensation changed the actual substance, one could not identify water (the liquid) with the compound of hydrogen and oxygen. (These statements must be qualified by pointing out that even Lavoisier did not think that strictly water and water-vapour are the same species, for water-vapour is water combined with caloric).

What these remarks show is that before Lavoisier it was reasonable to hold that different phases did not always differentiate between substances (and that even after Lavoisier the matter was not fully settled until the end of the caloric theory). In which case ‘gaseous water’ could not simply be a contradiction in terms. Rather it was a synthetic question whether freezing and evaporation involved a change in species¹³. The fact that many chemists did regard such changes as changes in species is no objection to essentialism. Indeed, as Kripke shows, a large part of the motivation for the causal theory is that it permits false beliefs about the

¹²Note that Locke is working with an idiolect view of language — we each have our own personal languages that we coordinate publicly. Without that, one might wonder why the proper response to the Jamaican visitor would not be ‘you’ve misused the word ‘water’; part of its definition is that it is a liquid’ — if, of course, water does have only a nominal essence.

¹³For this point and much other useful information on the subject I am very grateful to Robin Hendry.

referents of the terms one uses. In the much-used fictional example, Lois Lane believes that Clark Kent and Superman are different people. But that is consistent with their being the same individual and necessarily so. Similarly someone who recalls Eric Blair from their shared schooldays at Eton might not know that he is the same man as George Orwell, the now famous author of *Down and Out in Paris and London*, and indeed might naturally think there are different individuals, in which case his belief is a necessary falsehood.

It is noteworthy that Kuhn accepts the causal and essentialist story when it comes to the use of property names. He is even willing to agree that it comes very close to working just as precisely for some kind terms, such as 'gold' [Kuhn 1990a, 309]. (It is worth asking whether Kuhn's partial acceptance of the causal-essentialist story is compatible with the Wittgensteinian interpretation given by Read and Sharrock). To defend his incommensurability thesis, what is required is that the story break down for key terms across scientific revolutions, even if it holds for others. Since water is a point of focus in the Chemical Revolution, Kuhn wants to show that there isn't continuity of reference of 'water' across that revolution, even though there may have been for 'gold' and other kind terms that did not play such a central role. However, given the partial acceptance of the causal-essentialist thesis, Kuhn needs to have an especially convincing reason as to why what holds in the case of names (and perhaps some kind terms) does not hold in the case of 'water'. For in the case of names, we see that 'N' and 'M' might be rigid designators referring to the same individual (hence necessarily $N=M$), yet someone might believe that N and M are not identical. And the same response is available to his objection based on the widespread (but not universal) view, before 1750, that steam and ice are not water. For the essentialist may (and should) respond that those involved had a false belief: 'steam', 'ice', 'water' all refer to the same substance and always have done. Those like Locke who thought otherwise were mistaken and those like his fictional Jamaican who thought that ice is hardened water were correct. The former would have been just as mistaken as someone who, to borrow Locke's example, thought that solid and liquid gold are different substances (which Locke acknowledges they are not). Kuhn has given no reason to reject this account of matters. As remarked above, the causal theory of reference does not guarantee continuity of reference. But it does make it more likely, even through revolutionary scientific changes. To make out a case for discontinuity of reference or extension Kuhn needed to give far more careful argument and detailed historical support than he did provide.

How damaging would it be were we to admit that ‘ x is water’ did analytically entail ‘ x is liquid’ in 1750? I have suggested that this would leave untouched the conclusion that all water_{1750} is H_2O even though it would not be correct to say that all substances that are composed of H_2O are water_{1750} (where water_{1750} is what was being referred to by ‘water’ when used in 1750) [Bird 2000, 183-184]. For we can form the expression ‘liquid water’ where our term ‘water’ covers all three phases. And thus if water_{1750} is necessarily liquid water and water is essentially H_2O , then water_{1750} is also essentially H_2O .

Kuhn complains that if we are required to use descriptive terms in the specification of reference then we are back with the Fregean position that the causal theory was designed to avoid. But the cases are importantly different. For if we supplement the rigid designation with descriptive terms, the essentialist conclusions still follow. In Putnam’s arguments the problem was the insufficiency of a purely descriptive semantics for kind terms; his conclusions are thus consistent with a mixed causal-descriptive semantics (an approach which for many philosophers has supplanted the pure causal view while retaining its insights).

What is important in all this is whether the incommensurabilist threat to realism survives. Kuhn has not given us a reason to suppose that it does. Indeed it is questionable whether the incommensurability thesis itself survives. I have suggested that Kuhn’s arguments against the view that our term ‘water’ is the same as or translates the mid-eighteenth century term ‘water’ (in the mouths of scientists) are very weak (even less does he show that the terms have different references). But note that in the course of his arguments that there is no such equivalence Kuhn supplies an alternative translation: ‘water’ (as used in 1750) is translated by ‘liquid water’ (as used by today’s scientists) and has the same reference as ‘liquid H_2O ’. He concedes that modern science is capable of picking out the stuff that people in 1750 labelled ‘water’ [Kuhn 1990a, 312]. If so, then the realists have what they want, a way of making clear that what scientists today say is in contradiction to some of what scientists in 1750 were saying. They thus are able to explain how, if today’s scientists are right, science has corrected and added to the science of yesteryear.

9 Superficial properties are essential?

Kuhn regards the following argument as significant.

The so-called superficial properties are no less necessary than their apparent essential successors. To say that water is liquid H_2O is to locate it within an elaborate lexical and theoretical system. Given that system, as one must be in order to use the label, one can in principle predict the superficial properties of water (just as one could of those of XYZ), compute its boiling and freezing points, the optical wavelengths that it will transmit, and so on. If water is liquid H_2O , then these properties are necessary to it. If they were not realized in practice, that would be a reason to doubt that water really was H_2O [Kuhn 1990, 312-313].

Again Kuhn overstates the scientific and historical case. The theoretical framework required to state that water is (liquid) H_2O was in place by the time of Berzelius in the mid-nineteenth century. But that framework was not sufficient to permit Berzelius to make the calculations Kuhn refers to, even in principle. The theoretical framework required to be able to assert that water is (liquid) H_2O is much thinner than Kuhn supposes.

But Kuhn's remarks are in any case once again irrelevant. The superficial properties would be necessary if the laws of nature contained within the relevant theories and used to compute the details of boiling point, freezing point, and so on, are themselves necessary. But if the laws are contingent, then it is much more difficult to show that the superficial properties are necessary. Most philosophers of science have held the laws of nature to be contingent and have taken it that the superficial properties are correspondingly also contingent.

Now, it is possible for contingent laws to support necessary superficial properties. But the arguments are subtle and highly contentious, and have only recently been explored [Bird 2001], [Bird 2002]. Kuhn gives us no reason to think that with contingent laws the superficial properties will be anything but contingent too. Some philosophers do think that the laws of nature will be necessary. In which case, given the necessity of $\text{water} = H_2O$ the superficial properties will be necessary also. Again this view is contentious and Kuhn hasn't said anything relevant to it. Finally, even if the superficial properties are necessary, as Kuhn maintains, that doesn't show that they are essential. For as Kit Fine reminds us, the essence of an entity is much more restricted than the set of its necessary properties [Fine 1994]. We may thus distinguish between the essence of a substance and the non-essential but nonetheless necessary properties that follow from its essence.

10 Conclusion

We started by considering the apparent threat to scientific realism posed by Kuhn's incommensurability thesis. The common response by the realist is to appeal to referentialism. Referentialism shows how incommensurability as a thesis about the mutual untranslatability of theories emanating from different paradigms is after all consistent with scientific realism. The referentialist response has been elaborated in particular in the form of the causal theory of reference. Related to the causal theory (but by no means identical to it) is the thesis of essentialism.

Essentialism and the causal theory of reference have come in for considerable criticism, and the arguments are still live today. I have not discussed let alone deflected any of these criticisms here. In this paper I have considered only the criticisms brought to bear by Kuhn and elaborated by Read and Sharrock (which are very different from the criticisms usually discussed in the literature). Do their arguments provide some distinctive reason to doubt the causal theory of reference and essentialism, and thereby undermine the referentialist defence of scientific realism against the threat posed by the incommensurability thesis? Did Kuhn spot something that other philosophers have missed?

Despite the best efforts of Read and Sharrock to resurrect Kuhn's criticisms, the fact is that Kuhn did not fully understand Putnam's argument and his criticisms are largely irrelevant. Those criticisms do not give the scientific realist a reason to doubt her views. Of course there are other reasons that have been advanced for doubting scientific realism and these have not been under consideration here. Furthermore, Kuhn's own approach is itself an anti-realist one¹⁴. This creates difficulty in assessing Kuhn's criticisms of referentialism. For in so far as referentialism is a component of scientific realism, a criticism that assumes an anti-realist viewpoint is question-begging. For that reason, I have not considered Kuhnian ways of adding to the points he makes, that together present an alternative, inherently anti-realist viewpoint. For example, Putnam's arguments are framed in such a way that assumes that scientists can get to know the way things really are. While I do not think that the arguments need to be framed in an overtly realist manner, I have not here responded to criticisms of a Kuhnian sort that object to the realist way of framing the argument. For what we are considering here is the refer-

¹⁴Indeed, on some ways of understanding Kuhn's incommensurability thesis, the argument is from anti-realism to incommensurability rather than the other way around. For a discussion of the relationship between anti-realism and incommensurability see [Bird 2003].

entialist defence of scientific realism against an alleged criticism. And it is clearly legitimate for the realist to frame her defensive arguments in realist terms.

Given this dialectic, it needs to be repeated that the primary aim of the paper is limited to defending the referentialist version of realism against a particular set of criticisms, and not to defending realism against other kinds of criticism, let alone establishing the superiority of realism to its competitors. The primary aim is achieved so long as referential continuity remains an open possibility at the end of this paper. I have not sought to show that there definitely is referential continuity across revolutions in key cases. This would require more detailed historical work than I can present here, although I do suggest that the historical details presented in the case of 'water' point in this direction. That additional work would be required to show the superiority of the realist viewpoint, and that, as I have said, is not my current aim. The latter is limited to showing how referentialism can support realism and defending the causal version of referentialism against Kuhn's specific criticisms. The conclusion I reach is that the realist can mount a fully adequate defence.