

# A Critique of Putnam's Principle of the Benefit of Doubt

Igor Douven

Department of Philosophy, Utrecht University

E-mail: igor.douven@phil.uu.nl

## Abstract

The Principle of the Benefit of Doubt dictates that, whenever reasonably possible, we interpret earlier-day scientists as referring to entities posited by current science. Putnam has presented the principle as supplementary to his Causal Theory of Reference in order to make this theory generally applicable to scientific terms. The present paper argues that the principle is of doubtful standing. In particular, it will be argued that the principle lacks a justification and, indeed, is unjustifiable as it stands.

## 1 INTRODUCTION

In the 1970s, Putnam, in a number of influential papers, developed an account of meaning, reference, and their interrelations, which radically departed from the views then current in semantics.<sup>1</sup> Putnam's theory builds on Kripke's theory of naming, a theory that gives an intuitively plausible account of the naming of words for observable objects and properties but that does not immediately generalize to theoretical terms. In order to make it applicable to such terms, Putnam invokes what he calls the Principle of the Benefit of Doubt (hereafter PBD). Roughly, PBD dictates that, whenever reasonably possible, we interpret earlier-day scientists as referring to entities posited by current science. Although Putnam's theory has been criticized for various reasons,<sup>2</sup> PBD has received relatively little critical attention. This is surprising, given the principle's crucial role in Putnam's theory and given that, as I will argue, the principle lacks any justification. In effect, I will argue that the principle, as it stands, is *unjustifiable*, for as it stands it is incoherent, and even if it were not it is, barring further supplementation, unequal to its purported task.

The paper is organized as follows. In section 2 the views on meaning and reference Putnam challenged as well as Putnam's alternative proposal will be

---

<sup>1</sup>Cf. Putnam (1975a), (1975b), (1975c). Putnam's students Boyd and Devitt have also contributed significantly to the development of the theory; see Boyd (1979), Devitt (1981).

<sup>2</sup>Most importantly, critics have opposed against the metaphysics of natural kinds and the 'non-Humean' variety of causality that both seem to be presupposed by Putnam's theory. A 'later' Putnam has very actively taken part in criticizing these presuppositions; see for instance his (1990a), (1990b), (1992), (1994). For other criticisms see e.g. Lepore and Loewer (1988), van Brakel (1992), Douven and van Brakel (1998).

described. In the same section PBD and the motivation for it in the context of Putnam's theory will be considered in more detail. In section 3 I consider the question of PBD's justification and discuss and dismiss the only candidate justification Putnam suggests for it. In section 4 I then argue that PBD is incoherent. There we will also have a look at a principle that may serve the same purpose as PBD but that does not suffer from the same defect. In section 5 I argue that, even if PBD should be coherent, the principle stands in need of further supplementation lest it comes down to an implicit endorsement (at least for theoretical terms) of the very type of account of meaning and reference Putnam's was meant to replace; the same will be seen to hold for the alternative principle discussed in section 4.

## 2 INCOMMENSURABILITY AND THE CAUSAL THEORY OF REFERENCE

Holistic or network theories of meaning identify a term's meaning with its place in the theory in which it occurs, or, less metaphorically, with how the definitions and laws of the theory relate it to the other terms occurring in the same theory (cf. for instance Feyerabend 1962:82, 1965:180). Consequently, on such accounts, there can be no change in theory without a change in the meanings of its terms. This, in combination with the Fregean doctrine that meaning determines reference, has almost generally been taken to entail that no term in any theory can share its meaning or reference with a term in a distinct theory.<sup>3</sup>

This conclusion, also called the thesis of the incommensurability of theories,<sup>4</sup> constitutes a major threat to the doctrine of convergent scientific realism. After all, a minimal requirement for the convergence of scientific knowledge is that successive theories can be about the same entities. One way realists have sought to divert the threat is by challenging and supplanting the views on meaning and reference that seem to underlie the incommensurability thesis.

Putnam's new theory, now commonly called the Causal Theory of Reference (CTR), undoubtedly is the best known attempt to rebut this thesis. It tries to avoid the anti-realist consequence of the older semantics by arguing that rather than that meaning determines reference, almost the opposite is true; a word's reference, on this view, is *part* of its meaning, and thereby to an important extent determines its meaning. The other important component of word meaning is what Putnam calls the stereotype, basically a set of beliefs held about a word's extension in a linguistic community.<sup>5</sup> Quite evidently, Putnam's account, while

---

<sup>3</sup>The argument may not be airtight. It has been challenged by Scheffler (1967) (see also Papineau 1987); Niiniluoto (1997) has recently defended it against Scheffler's critique. At any rate, I will here assume that meaning holism together with the Fregean doctrine leads to at least some (from a realist perspective unacceptable) sort of incommensurability.

<sup>4</sup>Or that, in any event, is how Putnam says to understand the notion of incommensurability (see his 1978:23, 1981:114). Sankey (1990) may well be right that at least Kuhn would not agree with this conception of incommensurability. Here I will not try to adjudicate on this matter and only consider whether Putnam's new semantics offers a solution to the incommensurability problem as Putnam himself conceives of it.

<sup>5</sup>Further components of meaning on Putnam's theory are a syntactic and a semantic marker. See Putnam (1975a) for a detailed exposition of the theory, or Hacking (1983) for a survey of

it countenances that theory change induces meaning shifts, does not preclude referential continuity between successive theories.

Putnam relied on Kripke's (1980) theory of naming to explain how a word gets associated with its reference (the Fregean answer—'through its meaning'—no longer being available, of course). Briefly, Kripke's idea is that a proper name, or a name for a natural kind, has its reference in virtue of a causal chain grounded in something very much like a baptismal ceremony. Typically, a thing or substance is pointed to and given a name, which is then passed on in the language community; later users of the word refer with the word to the thing or substance baptized, provided the word has been appropriately passed on to them. As Kripke (1980:94–97, 106f) points out, reference can also be fixed by providing a description of the referent rather than by ostension. He emphasizes, though, that such a description of the referent of a word should not be equated with the word's meaning.

Since many of the entities science is concerned with are not perceptually available in the way gold, cats and yellow things (to mention Kripke's favorite examples of natural kinds) are, it seems that, for most scientific terms, we have to resort to the latter method of introductory descriptions in order to fix their reference (cf. Lepplin 1979, Devitt and Sterelny 1987). It is widely acknowledged that fixing the reference of a scientific term '*t*' requires more than to describe *ts* as 'whatever causes such-and-such observable effect(s)'. As Niiniluoto (1997:448) points out, such indirect causal descriptions are, by themselves, too liberal; e.g., "[i]f phlogiston is simply defined to be the cause of combustion, then the term 'phlogiston' refers to oxygen". Though it is reasonable to assume that such indirect causal description will always be *part* of an introductory description, the latter must also contain "some initial description of some fundamental or important properties of [the thing or substance]" (Niiniluoto 1997:448f; cf. also Nola 1980, Kroon 1985).<sup>6</sup>

Now what are we to say if, from the standpoint of present-day science, the description once used to fix the reference of some scientific term still in use in one of our theories is *false*? For example, from our perspective both Mendel's description of the gene<sup>7</sup> and Bohr's early description of the electron are false. Should we therefore say that neither was really referring to anything? Of course scientific realists would like to maintain that Mendel was talking about genes—the very same entities as are postulated by modern biology—and that Bohr was talking about electrons—*our* electrons—notwithstanding the fact that nothing in the world satisfies their descriptions of these entities. It is thus a desideratum that by itself the falsity, from our perspective, of descriptions earlier-day scientists used to refer to unobservable entities does not preclude us to say that they were referring to entities posited by our science. At the same time it must remain possible to deny that Priestley was talking about phlogiston; Priestley did not

---

it.

<sup>6</sup>This does not, as Niiniluoto (1997:449) apparently thinks, drive us back to a Fregean descriptive theory of reference. That is only so if the introduced term is taken to be synonymous with the description used to specify its referent, which, as remarked earlier, one need not do.

<sup>7</sup>Mendel in fact never used the word 'gene' but talked instead about discrete particles as constituting the inherited 'factor'. It was the Danish biologist Wilhelm Johannsen who in 1909 coined 'gene' as a word for the Mendelian inherited factor.

just give a wrong description of phlogiston, there *is* no phlogiston in this world (or so we believe). A supplementation of the theory thus seems called for that can discriminate between, on the one hand, false term-introducing descriptions of the kind Mendel gave for ‘gene’ and Bohr for ‘electron’ and, on the other hand, false term-introducing descriptions of the kind Priestley gave for ‘phlogiston’.

Putnam’s Principle of the Benefit of Doubt is introduced for this purpose. The principle dictates that

when speakers specify a referent for a term they use by a description and, because of mistaken factual beliefs that these speakers have, that description fails to refer, we should assume that they would accept reasonable reformulations of their description (in cases where it is clear, given our knowledge, how their description should be reformulated so as to refer, and there is no ambiguity about how to do it in the practical context). (Putnam 1978:23f)

Since such reasonable reformulation seems possible in Mendel’s case, but not in Priestley’s, we can say that the word ‘gene’ (or rather ‘factor’—cf. note 7) in Mendel’s theory referred to genes and the word ‘phlogiston’ in Priestley’s theory to nothing. Taken in combination, CTR and PBD thus seem to offer a solution to the problem of incommensurability.

The remainder of the paper will be devoted to scrutinizing PBD.

### 3 WHAT JUSTIFICATION IS THERE FOR THE PRINCIPLE?

A host of arguments have been given, by Putnam, Kripke and many others, to convince us that the causal account is intuitively much more plausible than the older accounts of meaning and reference that it is meant to replace. While these arguments are generally regarded as being among the most persuasive in modern philosophy, they all turn on examples exclusively involving proper names and words for observable natural kinds. It is not at all evident from these arguments that the theory also gives the right account of reference and meaning for theoretical terms.<sup>8</sup> Thus there remains the question as to what justifies application of the theory to theoretical terms and, in particular, what justifies the principle crucial in applying the theory to such terms.

Initially PBD may strike one as being quite plausible. Surely we do not expect any scientist to object to a *reasonable* reformulation of a description he or she used for the specification of the referent of some theoretical term? Note, however, that this cannot count as a justification for PBD. For what needs justification is the assumption that, if a former scientist can reasonably be assumed to accept a (from our perspective) correct reformulation of an introductory description, then that provides sufficient ground to conclude there is referential continuity between him/her and us. And it is certainly not obvious that the assumption’s antecedent

---

<sup>8</sup>If we accept a principle of semantic uniformity according to which there must be a single theory of meaning and reference for the entire vocabulary of a language, then it is arguable that, if the causal theory correctly accounts for the meaning and reference of words for observable things and properties, it must also be the correct account of reference and meaning for theoretical terms. However, I have argued elsewhere (Douven 1998) that such principles of semantic uniformity, although widely accepted in modern semantics, are themselves ill-justified.

has anything to do with its consequent. Putnam seems to be aware that PBD needs some justification of its own (i.e., one independent of the justification provided for CTR) and in his (1978) he does propose such a justification for the principle. What he claims there is that PBD is justified in virtue of its (alleged) ability to block the so-called pessimistic meta-induction. It is easily seen, however, that if this (alleged) justification should be the only one for PBD—and to the best of my knowledge it is the only one proffered in the literature—then application of PBD lacks any warrant.

Putnam (1978:24f) says:

[W]hat if we accept a theory from the standpoint of which electrons are like phlogiston? Then we will have to say electrons don't really exist. What if this keeps happening? What if *all* the theoretical entities postulated by one generation (molecules, genes, etc., as well as electrons) invariably 'don't exist' from the standpoint of later science? This is, of course, a form of the old sceptical 'argument from error'—how do you know you aren't in error *now*? . . . One reason this is a serious worry is that eventually the following meta-induction becomes overwhelmingly compelling: *just as no term used in the science of more than fifty (or whatever) years ago referred, so it will turn out that no term used now . . . refers.*

This argument is now commonly referred to as the 'pessimistic meta-induction'; scientific realists regard it as one of the major weapons in the hands of their opponents.<sup>9</sup>

It might seem that, although it threatens convergent scientific realism, the pessimistic meta-induction does not, or at least not necessarily, pose a problem for Putnam's new account of meaning and reference. That is not so, however. Since it makes meaning and reference so crucially dependent on actual causal relations between words and things in the world, Putnam's theory seems unable to explain how we understand the meanings of such 'non-entity terms' as 'phlogiston' or 'caloric' (terms still current today, albeit mainly in history and philosophy of science seminars), as most realists think we do. There can, for instance, be no causal chain linking our present use of the term 'phlogiston' to the substance phlogiston, for there is no, and there never was, such a substance. At best, such cases show that Putnam's theory cannot be complete. (At worst, they show causal connections are irrelevant to what our words mean and hence that Putnam's theory is wrong.) But, of course, if *all* terms were like 'phlogiston', as the meta-induction would have us conclude, then Putnam's theory would not just be incomplete but would be entirely useless.<sup>10</sup> Thus Putnam (1978:25) rightly remarks: "It must obviously be a desideratum for the theory of reference that this meta-induction be blocked". The ensuing remark—"[T]hat is one justification for the principle of the benefit of doubt"—is deeply puzzling, however, and that for at least two reasons.

First, Putnam thinks it would be unreasonable to accord Priestley the benefit of the doubt and to say 'phlogiston' referred (*ibid.*). Now, one can easily formulate a pessimistic meta-induction only involving terms none of which can reasonably

---

<sup>9</sup>For a much more detailed version of the argument see Laudan (1981).

<sup>10</sup>The relationship between the pessimistic meta-induction and Putnam's account of meaning and reference is discussed in much greater detail in Douven and van Brakel (1995).

be regarded as referring to some entity postulated by present-day science. In Douven and van Brakel (1995:7) the following example of such a meta-induction is given:

Just as according to Science<sub>1800</sub> the term ‘phlogiston’ had no referent, just as according to Science<sub>1850</sub> ‘phlogiston’ and ‘caloric’ had no referent, just as according to Science<sub>1900</sub> ‘phlogiston’, ‘caloric’ and ‘ether’ had no referent ... (etcetera) so, according to Science<sub>2050</sub>, ‘phlogiston’, ‘caloric’, ‘ether’, ... , ... and ‘electron’ (or any other term from current science) will have no referent, and, according to Science<sub>2100</sub>, ‘phlogiston’, ... , ... , ‘electron’ and ‘gene’ will have no referent, and, according to Science<sub>2150</sub>, ... and ... will have no referent ... (etcetera). And so will, finally, according to some future theory, all or almost all of the terms from current science appear to have no referent.

It is pretty clear that against this argument PBD is of no help.<sup>11</sup>

Secondly, suppose the induction were exclusively based on evidence involving ‘entities’ once but no longer thought to exist but which can be said to be in important respects similar to entities postulated by present-day science (e.g. Rutherford’s ‘electrons-with-trajectories’). In that case it seems arguable that PBD *can* help reduce or even take away the meta-induction’s skeptical force. It should be noticed, however, that our present concern is finding an adequate *justification* for PBD. And it seems to me that the mere fact that a principle can help defending a certain philosophical position does not by itself constitute a reason for finding that principle acceptable.

But perhaps I’m wrong; perhaps the main justification for PBD, and for principles like it, *is* to be sought in what they accomplish. It seems arguable that, quite apart from its threatening scientific realism, the incommensurability thesis is problematical because it seems just too much to believe that any theory change, however minimal, induces referential shifts. So then it may seem to some that if PBD is able to help refute that thesis that already provides sufficient justification for it. However, there is no need to squabble about this, for I do not think PBD *is* of any help against the incommensurability problem. In fact, as will be seen in the next section, the principle as formulated by Putnam is not even coherent.

#### 4 PBD IS INCOHERENT

From some of Putnam’s remarks concerning PBD, one can easily get the impression that it is *only* a methodological maxim governing our practice of interpreting former scientists. For instance, Putnam (1978) says that according to this principle “we should *assume* that [speakers] would accept reasonable reformulations of [their term-introducing descriptions]” (23f; my italics), that we must “*treat* Bohr as referring to these particles [i.e. electrons]” (24; my italics) or again that “we should ... *take* Bohr to have been referring to what we call ‘electrons’” (*ibid.*;

---

<sup>11</sup>Of course, one may doubt the force of such meta-inductions, whatever their precise formulation. However, if these arguments are non-starters, as some authors maintain (cf. for instance Devitt 1991, Cummiskey 1992, Leplin 1997), they already for that reason cannot provide a justification for PBD.

my italics). PBD might thus seem to imply no more than the following condition under which terms occurring in successive theories should be interpreted as having the same reference:

PBD' A term '*t*' in theory *T* should be interpreted as having the same reference as '*t*<sup>\*</sup>' in a successor theory *T*<sup>\*</sup> of *T* (with possibly '*t*' = '*t*<sup>\*</sup>') exactly if the description *T*'s proponents used to specify the referent of '*t*' can, from the perspective of *T*<sup>\*</sup>, be reasonably reformulated so as to make '*t*' refer to the same thing(s) '*t*<sup>\*</sup>' refers to.

But if all that PBD says has to do with how we should *interpret* others, implying nothing about whether there *is* any shared reference, then it is hard to see how it could be of any help in answering the incommensurability thesis. For to do this, realists must establish that, in general, later theories *are* advances on truth as compared to their predecessors, not that they can be *interpreted* or *taken* or *treated* as advances on truth; the latter conclusion would just be too weak to be any good.

So I take it that PBD is not just meant to state the condition under which some term occurring in a now discarded theory can be interpreted as having the same reference as a term currently employed, but that it also intends to state the following condition under which such terms actually have a common reference:

PBD'' A term '*t*' in theory *T* has the same reference as '*t*<sup>\*</sup>' in a successor theory *T*<sup>\*</sup> of *T* (with possibly '*t*' = '*t*<sup>\*</sup>') exactly if the description used by *T*'s proponents to specify the reference of '*t*' can, from the perspective of *T*<sup>\*</sup>, be reasonably reformulated so as to make '*t*' refer to the same thing(s) '*t*<sup>\*</sup>' refers to.

But this condition is incoherent. The problem is that the relation between descriptions and perspectives from which a certain description can be reasonably reformulated is not transitive: it may be that from our perspective Bohr's description of the electron can be reasonably reformulated, and that our description of the electron can be reasonably reformulated from the perspective of a future generation of scientists, but that Bohr's description *cannot* be reasonably reformulated from that future perspective. By PBD'' this would imply that Bohr was talking about the same electrons as we do, that we are talking about the same electrons as (we suppose) future scientists will do, *and* that these future scientists will not be talking about the same electrons as Bohr was talking about. This shows that, since sameness of reference *is* transitive, application of the principle may lead to outright contradiction: the future scientists we imagined both are and are not talking about the same electrons as Bohr was talking about.

The problem could have been avoided, I think, by adopting instead of PBD the following alternative principle:

PBD\* A term '*t*' occurring in theory *T* refers to the entity or kind of entities *E* if the description used by *T*'s proponents to specify the reference of '*t*' is closer to the truth about *E* than it is to the truth about any other entity or kind of entities.

PBD\* is, in essence, the supplementary principle to CTR Niiniluoto (1997) proposes. He makes it much more precise with the help of the notions of approximate

truth and truthlikeness which in his work have very definite, technical meanings (cf. Niiniluoto 1987). There is for us no need to go into these intricacies; I here assume the idea of a description being closer to the truth about one thing than to the truth about another to be unproblematic. Niiniluoto also points out that the principle has to be qualified in that some threshold requirement should be added to PBD\*: The description should be *sufficiently* close to the truth about some entity for it to refer at all, where some specification of the sufficiency condition has to be provided; this too I will leave aside here.<sup>12</sup>

Like PBD, PBD\* allows terms in distinct theories to have the same reference. E.g., it may be that Bohr's and our description of the referent of the term 'electron', though different, are closer to the truth about one and the same kind of entity than to the truth about any other kind of entity. So it seems that the combination of CTR and PBD\* is sufficient to block incommensurability.

Since PBD\* is superficially very similar to PBD it may not be immediately clear why the new principle does not face the same problem as PBD. The crucial difference between the two principles is this: Whereas PBD is concerned with the question as to whether terms in different theories are and can be taken to be co-designative, PBD\* states a condition for a term to refer period. The latter is entirely independent of what can or cannot be reasonably reformulated from a certain perspective. Thus, if theory  $T^*$  is the successor of  $T$ , it may be possible from the perspective of  $T^*$  to reasonably reformulate the description for some term ' $t$ ' occurring in  $T$  so as to make it refer to what ' $t^*$ ' in  $T^*$  is supposed to refer to, and yet ' $t$ ' in  $T$  and ' $t^*$ ' in  $T^*$  may, according to PBD\*, both fail to refer. Conversely, it cannot be excluded that, according to PBD\*, both ' $t$ ' and ' $t^*$ ' refer to  $ts$ , and yet there is, from the perspective of  $T^*$  no reasonable reformulation of the  $T$ -dependent description associated with  $t$ . The question as to whether terms in distinct theories have the same reference thus has nothing to do with whether term-introducing descriptions given by proponents of the one theory are regarded as being sufficiently close to theirs by proponents of the other. Hence PBD\* cannot lead to the problem PBD was seen to give rise to.

## 5 PBD NEEDS FURTHER SUPPLEMENTATION

Initially one is certainly inclined to agree with Putnam (1978:24) that, according to our best current knowledge,

there are particles which approximately fit [the Bohr-Rutherford description of an electron]: they have the right charge, the right mass, and they

---

<sup>12</sup>Putnam's (1975c:275ff) discussion of Wilson's Principle of Charity (according to which "we should assign the designatum that makes the largest possible number of the speaker's beliefs true"—275) suggests that PBD\* may require qualification in another respect. There Putnam imagines a kind of entities he calls 'shmelectrons' which only exist in some remote region of our universe. He further imagines that these entities have all or most of the properties Bohr thought electrons had, while electrons do not. As Putnam points out, application of Charity would now oblige us to say Bohr was really talking about shmelectrons, which intuitively is certainly wrong. As it stands, PBD\* suffers the same defect. This is easily fixed however. One way to do this is by demanding that the indirect causal description part of the introductory description be *true* (and perhaps also that such causal descriptions be stated in the form: 'responsible for effects such as occurred then-and-then, there-and-there').

are responsible for key effects which Bohr-Rutherford explained in terms of 'electrons'.

For all we know, electrons do not, as Rutherford and Bohr (in 1910, say) thought, orbit the nucleus much in the same way as the Earth orbits the Sun. But it seems entirely reasonable to assume Rutherford and Bohr would accept reformulation of their description so as to make it correct from our perspective (as of course we know, Bohr later in his career actually did accept such a reformulation). If that is so, then PBD demands that we identify the reference of the term 'electron' as used by Rutherford and Bohr in the beginning of this century with the reference of the term as we now use it.

However, we may well be too quick with our application of PBD here. We find it appropriate to accord Bohr the benefit of the doubt in the case of the electron because we judge the description he associated with the term 'electron' as being sufficiently close to the description we associate with the term; both ascribe the same mass and charge to electrons and both agree on what observable effects are caused by electrons. But in saying that both descriptions to a large part ascribe the same properties to electrons we are obviously assuming that we can identify the reference of such terms as 'charge' and 'mass' occurring in Bohr's description with the reference these terms have as they occur in our theories. Otherwise put, in claiming that there are entities approximately fitting Bohr's 1910 description of the electron, Putnam must be assuming that the theoretical terms occurring in that description (apart from 'electron', perhaps) have retained their reference since 1910. Whether we can be reasonably confident that they have, despite various changes in theory since that time, is of course exactly the question at issue.

More formally put the problem is that, as Putnam presents the situation, it is as if we must consider whether it would be reasonable to interpret 'X' as referring to electrons in each of the sentences ' $\forall x(Xx \rightarrow Px)$ ', ' $\forall x(Xx \rightarrow Qx)$ ' and ' $\forall x(Xx \rightarrow Rx)$ '. If, from our current perspective, electrons have properties  $P$  and  $Q$ , though perhaps not  $R$ , it may be arguable that the interpretation *is* reasonable. However, as we just saw, in order not to beg any Kuhnian or Feyerabendian issues, we should represent the situation as one in which it must be decided whether the class of electrons could be reasonably assigned as reference to 'X' in the sentences ' $\forall x(Xx \rightarrow X'x)$ ', ' $\forall x(Xx \rightarrow X''x)$ ', etc., that is, in sentences in which the properties ascribed to whatever has the property  $X$  are themselves treated as 'blanks'.

Would it help if we were to apply PBD to *all* theoretical terms occurring in Bohr's 1910 theory? Identify the referents of Bohr's 'mass', 'charge', etc., with our mass, charge, and so on, and Bohr's description of the electron does come out close to ours.

While this is correct, it should be noticed that this procedure gives us something very much like the network theory of meaning back, at least for theoretical terms. Whether an assignment of a certain referent to one of Bohr's terms is reasonable or not depends on what referents we assign to the other terms in the network. Assignments of referents, and hence, on CTR, of meanings, to theoretical terms will thus have to be evaluated clusterwise. In other words, it is again a term's location in the theoretical network that determines its reference, and

hence again a term's place in the network that determines what it means. Surely the referents that thus get assigned are all entities acknowledged by current scientific theory. So more precisely put the theory would now be that which of the entities postulated by current science are assigned to the terms of some formerly accepted but now discarded theory depends on the network constituted by the latter. It is hard to see how that could make a relevant difference, however.

Even if it does, the suggested unrestricted application of PBD cannot be acceptable to a proponent of CTR. As was pointed out, PBD was supposed to be strict enough to not make, e.g., 'phlogiston' come out as a referring term. But when unrestricted application is admitted, it is not clear that this can be avoided. It presumably would not take much imagination to find some assignment of referents to the central terms of phlogiston-theory that would make the description Priestley used to specify the reference of 'phlogiston' come out close to one we use to specify the reference of some term currently in use. On the present proposal, this would be a legitimate practice. More specifically, on this proposal it is unclear what such qualifications as that PBD should be applied to a term only when a *reasonable* reformulation of its introductory description is possible, and that there should be no *ambiguity* about how to do it practically, are supposed to come to; the terms 'reasonable' and 'unambiguous' as applied to a possible reformulation of an introductory description would seem to have become entirely void.

This time the problem is not circumvented by adopting PBD\* instead of PBD. Recall that according to PBD\* a term refers to the entity or kind of entities that maximizes the 'truth-closeness' of its associated description. But for this to make sense the truths occurring in the description must themselves have their reference already fixed. Since such descriptions will typically contain themselves theoretical terms, this requires some account that explains how *these* theoretical terms have their reference fixed. And from our discussion above we know that it is no help to modify PBD\* and state that the terms of a theory refer to those (classes of) objects that, when collectively assigned as referents to the various terms, maximize the theory's truth-closeness.

The foregoing does not by itself show that there could not be any role for PBD/PBD\* in supplementing CTR, but merely that, to make the theory applicable to theoretical terms, some further supplementation is needed. If, for instance, we had some account that could explain how, in the example of Bohr's electron, at least 'mass' and 'charge' have retained their reference since 1910, and that does so without itself relying on PBD/PBD\*, we might want to invoke PBD/PBD\* to conclude to the referential continuity of the term 'electron' since 1910.<sup>13</sup> However, it seems unlikely that, if we really had a theory that explained how certain theoretical terms retain their reference through theory change, this theory would not work for all theoretical terms alike. In that case there would evidently be no call for an application of PBD/PBD\* any longer.

---

<sup>13</sup>It might be objected that the meaning of 'electron' in Bohr's theory in that case still would be dependent on those of other theoretical terms. However, this would not be meaning holism. Meaning holism is a general thesis, a thesis about what determines the meaning of any (theoretical) term. It is certainly compatible with CTR that *some* theoretical terms depend for their meaning on the meanings of other terms; cf. for instance Devitt (1990).

## REFERENCES

- Boyd, R. (1979) 'Metaphor and Theory Change', in: A. Ortony (ed.) *Metaphor and Thought*, Cambridge: Cambridge University Press, pp. 356–408.
- Cummiskey, D. (1992) 'Reference Failure and Scientific Realism: a Response to the Meta-Induction', *The British Journal for the Philosophy of Science* 43:21–40.
- Devitt, M. (1981) *Designation*, New York: Columbia University Press.
- Devitt, M. (1990) 'Meanings Just Ain't in the Head', in: G. Boolos (ed.) *Meaning and Method*, Cambridge: Cambridge University Press, pp. 79–104.
- Devitt, M. (1991) *Realism and Truth*, Oxford: Blackwell.
- Devitt, M. and K. Sterelny (1987) *Language and Reality: An Introduction to the Philosophy of Language*, Oxford: Blackwell.
- Douven, I. (1998) 'Truly Empiricist Semantics', *Dialectica* 52:127–151.
- Douven, I. and J. van Brakel (1995) 'Is Scientific Realism an Empirical Hypothesis?', *Dialectica* 49:3–14.
- Douven, I. and J. van Brakel (1998) 'Can the World Help Us in Fixing the Reference of Natural Kind Terms?', *Journal for General Philosophy of Science* 29:59–70.
- Feyerabend, P. (1962) 'Explanation, Reduction and Empiricism', in: H. Feigl and G. Maxwell (eds) *Minnesota Studies in the Philosophy of Science*, vol. 3, Minneapolis MI: University of Minnesota Press, pp. 28–97.
- Feyerabend, P. (1965) 'Problems of Empiricism', in: R. Colodny (ed.) *Beyond the Edge of Certainty*, Englewood Cliffs NJ: Prentice Hall, pp. 145–260.
- Hacking, I. (1983) *Representing and Intervening*, Cambridge: Cambridge University Press.
- Kripke, S. (1980) *Naming and Necessity*, Oxford: Blackwell.
- Kroon, F. (1985) 'Theoretical Terms and the Causal View of Reference', *Australasian Journal of Philosophy* 63:143–166.
- Leplin, J. (1979) 'Reference and Scientific Realism', *Studies in History and Philosophy of Science* 10:265–285.
- Laudan, L. (1981) 'A Confutation of Convergent Realism', *Philosophy of Science* 48:19–49.
- Leplin, J. (1997) *A Novel Defense of Scientific Realism*, Oxford: Oxford University Press.
- Lepore, E. and B. Loewer (1988) 'A Putnam's Progress', in: P. French, Th. Uehling and H. Wettstein (eds) *Midwest Studies in Philosophy*, vol. 12, Minneapolis: University of Minnesota Press, pp. 459–473.

- Niiniluoto, I. (1987) *Truthlikeness*, Dordrecht: Reidel.
- Niiniluoto, I. (1997) 'Reference Invariance and Truthlikeness', *Philosophy of Science* 64:546–554.
- Nola, R. (1980) 'Fixing the Reference of Theoretical Terms', *Philosophy of Science* 47:505–531.
- Papineau, D. (1987) *Reality and Representation*, Oxford: Blackwell.
- Putnam, H. (1975) *Mind, Language and Reality: Philosophical Papers, vol. 2*, Cambridge: Cambridge University Press.
- Putnam, H. (1975a) 'Explanation and Reference', in his (1975), pp. 196–214.
- Putnam, H. (1975b) 'The Meaning of 'Meaning'', in his (1975), pp. 215–271.
- Putnam, H. (1975c) 'Language and Reality', in his (1975), pp. 272–290.
- Putnam, H. (1978) *Meaning and the Moral Sciences*, London: Routledge and Kegan Paul.
- Putnam, H. (1981) *Reason, Truth, and History*, Cambridge: Cambridge University Press.
- Putnam, H. (1990) *Realism with a Human Face*, Cambridge MA: Harvard University Press.
- Putnam, H. (1990a) 'Is Water Necessarily H<sub>2</sub>O?', in his (1990), pp. 54–79.
- Putnam, H. (1990b) 'Is the Causal Structure of the Physical Itself Something Physical?', in his (1990), pp. 80–95.
- Putnam, H. (1992) *Renewing Philosophy*, Cambridge MA: Harvard University Press.
- Putnam, H. (1994) 'Model Theory and the "Factuality" of Semantics', in his *Words and Life*, Cambridge MA: Harvard University Press, pp. 351–375.
- Sankey, H. (1990) 'In Defence of Untranslatability', *Australasian Journal of Philosophy* 68:1–21.
- Scheffler, I. (1967) *Science and Subjectivity*, Indianapolis: Bobbs-Merill.
- van Brakel, J. (1992) 'Natural Kinds and Manifest Forms of Life', *Dialectica* 46:243–263.