



OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

The British Society for the Philosophy of Science

Reference Failure and Scientific Realism: A Response to the Meta-Induction

Author(s): D. Cummiskey

Reviewed work(s):

Source: *The British Journal for the Philosophy of Science*, Vol. 43, No. 1 (Mar., 1992), pp. 21-40

Published by: [Oxford University Press](#) on behalf of [The British Society for the Philosophy of Science](#)

Stable URL: <http://www.jstor.org/stable/687883>

Accessed: 05/12/2011 11:17

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Oxford University Press and *The British Society for the Philosophy of Science* are collaborating with JSTOR to digitize, preserve and extend access to *The British Journal for the Philosophy of Science*.

<http://www.jstor.org>

Reference Failure and Scientific Realism: a Response to the Meta-induction

D. CUMMISKEY

ABSTRACT

Pure causal theories of reference cannot account for cases of theoretical term reference failure and do not capture the scientific point of introducing new theoretical terminology. In order to account for paradigm cases of reference failure and the point of new theoretical terminology, a descriptive element must play a role in fixing the reference of theoretical terms. Richard Boyd's concept of theory constitutive metaphors provides the necessary descriptive element in reference fixing. In addition to providing a plausible account of reference failure and success, a metaphor approach to reference fixing provides the basis for a plausible realist account of the progress of science. Indeed, the metaphor approach undermines the sceptical force of the meta-induction and Laudan's objections to scientific realism.

1 *Introduction*

2 *Causal Theories of Reference and the Triviality of Scientific Realism*

3 *Reference Failure and Theory Constitute Metaphor*

4 *Scientific Realism and the Meta-induction*

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 (except maybe observation terms, if there are such) refers*. It must obviously be a desideratum for the theory of reference that this meta-induction be blocked (Hilary Putnam [1978] *Meaning and the Moral Sciences*, p. 25).

I INTRODUCTION¹

Realism about scientific entities and a theory of reference for scientific terms present two related problems. The first problem is whether causal theories of

¹ I am indebted to Joe Mendola, Peter Railton, and Larry Sklar for helpful discussions about

reference can adequately account for cases of reference failure for theoretical terms.² The second problem is the above meta-induction. I shall argue that only a hybrid causal and description theory of reference can adequately account for cases of reference failure for theoretical terms. In addition, we shall see that the theory of reference presented provides a plausible response to the meta-induction.

The hybrid theory is derived from Richard Boyd's article 'Metaphor and Theory Change.'³ Boyd argues that some metaphors, which he calls 'theory constitutive metaphors,' play a central role in the development of some scientific theories.⁴ Basically, metaphor provides a sense which limits the ways we (the linguistic community using the term) can be wrong about the entity or property in question. I maintain that it is only through the use of such metaphors that we are able to lock many theoretical terms on to some *particular* aspect of reality. As a first approximation, when scientists discover, through future research, that there is some aspect of reality which prompted the use of the term that is (roughly) similar to what the metaphor implies, then we have successfully referred. If, however, scientists are unsuccessful in finding any aspect of reality (roughly) similar to what the metaphor implies, then we fail to refer. The term is empty.

As we shall see in Section 2, some descriptive element must play a role in fixing the reference of an important class of theoretical terms. It would be foolish, however, to simply abandon the insights of the causal approach. If possible, the descriptive element should be compatible with the strengths of the pure causal theory. In addition, the account of reference fixing should dovetail with the explanatory point of introducing new scientific terms. Typically, new

various realisms. My thanks to Michael Devitt and Jim Joyce for comments on an early draft of this paper and to Keith Quillen, John Nolt, and Roger Jones for their comments on a recent draft. Finally, the careful comments, critical and constructive, of an anonymous referee have helped eliminate several errors and confusions and have prompted clearer responses to several important objections.

² For discussions of causal theories of reference see, for example, articles by Kripke, Putnam, Donnellan, and Gareth Evans in Schwartz (ed.) *Naming, Necessity, and Natural Kinds* (Cornell University Press, 1977). Also see Kripke *Naming and Necessity* (Harvard University Press, 1980); Putnam *Mind, Language, and Reality; Philosophical papers vol. 2*, chapters 11 and 12 (Princeton University Press, 1977); and Dummett *Frege: Philosophy of Language*, chapter 5 (Harvard University Press, 1973).

³ Boyd 'Metaphor and Theory Change' in *Metaphor and Thought*, Andrew Ortony (ed.) (Cambridge University Press, 1979). As we shall see, the argument developed in this article is different from Boyd's argument. Most prominently, (i) I do not make any use of 'epistemic access' which is a focal point of his paper; (ii) Boyd does not discuss the adequacy of causal theories of reference for theoretical terms and scientific realism; (iii) Boyd does not make use of Michael Devitt's 'quasi-perception' as a causal link to the world; (iv) I claim that the account of reference and scientific progress here presented disarms the meta-induction and undermines the more detailed challenge to realism presented by Larry Laudan in 'A Confutation of Convergent Realism,' (*Philosophy of Science*, 48, [1981], pp. 19–49). I am thus unsure how congenial Boyd would find the argument that follows. Nonetheless, my debt to Boyd's work is obvious.

⁴ For a detailed account of the differences between literary metaphors and theory constitutive metaphors see Boyd ([1979], pp. 356–64).

terminology is introduced in the context of a research project which is believed to have discovered new causally significant features of the world. Theory constitutive metaphors provide the necessary additional descriptive element and also capture the programmatic nature of scientific research.

The advantage of the metaphor approach is the flexibility of the descriptive element. The *flexibility* of metaphors provides an account of reference fixing which avoids the objections to typical description theories. Nonetheless, metaphor is rigid enough to provide a descriptive *constraint* which links the term to some specific aspect of reality. By providing flexible constraints, the metaphor approach best satisfies the requirement for a descriptive element in reference fixing. It is an additional, but equally important, virtue of the metaphor approach that the resulting explanation of reference failure (and success) provides an appealing picture of the progress of science and a plausible response to what Putnam has called the meta-induction.

In the philosophy of science, the debate over the adequacy of the meta-induction is primarily between realists and non-realist instrumentalists. This particular debate is distinct from the debate in semantic theory over the adequacy of 'metaphysical realism'. Briefly stated, even if one is a metaphysical anti-realist it is still an open question as to whether scientific realism or instrumentalism is the best *theory* of theoretical entities. As Putnam has argued, the truth of the empirical theory of scientific realism does not help us in solving the semantic problem 'as to how the *whole* representation, including the empirical theory of knowledge that is part of it, can determinately refer' (Putnam [1980], p. 477). Putnam calls himself an '*internal realist*' because he believes that all judgments about the truth of empirical theories are perspectively rooted and are thus just more empirical theory. Putnam calls himself an '*internal realist*' because of his views about scientific realism. The semantic realist debate leaves open the question of whether internal realism or internal instrumentalism is the correct view. In short, one can discuss the adequacy of scientific realism without making a commitment to metaphysical realism or internal realism. I make no commitment to realism as a semantic thesis. The meta-induction asserts that the *theory* that scientific theories are converging on an accurate representation of the world they study is inductively unwarranted. It is to this empirical induction that I respond.⁵

In his scientific realist mode Putnam, following Boyd, has argued that when understood as an 'over-arching empirical hypothesis', scientific realism involves 'two principles: (1) terms of a mature science typically *refer*, (2) the

⁵ On realism as a semantic issue see Dummett 'What is a Theory of Meaning? (II)' (in *Truth and Meaning*, Evans and McDowell (ed.), Clarendon Press, Oxford, 1978), and *Truth and Other Enigmas* (Harvard University Press, 1978); Putnam *Meaning and the Moral Sciences* (Routledge and Kegan Paul, 1978), and 'Models and Reality' (*Journal of Symbolic Logic* [1980], pp. 464–82). For a critical discussion of internal realism see Hacking *Representing and Intervening*; pp. 92–111 (Cambridge, 1983).

laws of a theory belonging to a mature science are typically approximately true.⁶ Actually, as Nancy Cartwright and Ian Hacking have pointed out, we may distinguish two distinct kinds of scientific realism: entity-realism and theory-realism.⁷ My concern is reference and entity-realism.⁸ I leave aside difficult questions about the concept of approximate truth and theory-realism.⁹

Recent arguments for realism argue by abduction (or inference to the best explanation) from the success of modern science to the truth of realism. There are two types of these arguments. The simple type argues from the individual successes of modern science to realism as the best explanation for these successes. The more complex type appeals to the methodology employed by modern science and argues that realism is the best explanation for the success of modern scientific methodology. In addition to developing the concept of theory constitutive metaphors, Boyd has developed these subtle *methodological* arguments for realism. It is thus worth noting that the arguments which follow are in no way connected to Boyd's methodological arguments for realism.

I am interested, however, in the most popular objection to the *simple* argument from success to realism. Arguments of the simple type for entity-realism allegedly fail because of the successes of past theories whose central terms we now believe failed to refer. Larry Laudan's 'A Confutation of Convergent Realism' is the most detailed and impressive objection to the simple argument for entity-realism.¹⁰ In addition to responding to the meta-induction, I maintain that Laudan's objection to the simple argument for entity-realism is hasty and inconclusive.¹¹ Thus, although I do not present any

⁶ Putnam [1978], p. 20.

⁷ Cartwright *How the Laws of Physics Lie* (Oxford, 1983) and Hacking 'Experimentation and Scientific Realism' (in *Scientific Realism* Jarrett Leplin (ed.) University of California Press, 1984). For brief arguments that these are distinct see Cartwright pp. 4–10, Hacking pp. 155–7 and Hardin and Rosenberg 'In Defence of Convergent Realism' pp. 606–8 (*Philosophy of Science*, 49 [1982], pp. 604–15).

⁸ I will not discuss Stephen Leed's argument, in 'Reference and Truth' (*Erkenntnis* [1973]), and will assume that a term genuinely refers only if the entity named by the term exists—'A' refers only if A exists. For an interesting discussion of Leed's argument see Michael Devitt *Realism and Truth* (Princeton [1984]).

⁹ For problems about the concept of approximate truth and theory-realism see Cartwright [1983]; Fine 'The Natural Ontological Attitude', pp. 89–91 (in *Scientific Realism* Jarrett Leplin (ed.), University of California Press, 1984); Laudan [1981]; and Larry Sklar 'Do Unborn Hypotheses Have Rights?' (*Pacific Philosophical Quarterly* [1981], pp. 17–29). On the other hand, for balance, also see Ernan McMullin's 'A Case for Scientific Realism,' in *Scientific Realism* Jarrett Leplin (ed.) (University of California Press, 1984); pp. 26ff. Cartwright's distinction between theoretical laws and phenomenological laws is also relevant ([1983], 1–4). The arguments of this paper require at most a phenomenological theory-realism. I do not, however, directly discuss these issues.

¹⁰ Laudan, 'A Confutation of Convergent Realism' (*Philosophy of Science*, 48 [1981], pp. 19–49).

¹¹ See Boyd 'The Current Status of Scientific Realism' (in *Scientific Realism* Jarrett Leplin (ed.) University of California Press, 1984) for the methodological argument for scientific realism. For the most interesting critical discussion of Boyd's arguments see Arthur Fine (in Leplin [1984]).

arguments for scientific realism, the arguments that follow purport to undermine the most widely accepted objection to scientific realism.

2 CAUSAL THEORIES OF REFERENCE AND THE TRIVIALITY OF SCIENTIFIC REALISM

Causal theories of reference fail to provide an account of reference failure for theoretical terms and thereby trivialize the referential claim of scientific realism. An adequate realist theory of reference must avoid this consequence.

The scientific realist claims that the development of science reveals progress in the ‘accommodation of language to the causal structure of the world.’ This involves, to quote Boyd,

the task of introducing terminology, and modified usage of existing terminology, so that linguistic categories are available which describe the causally and explanatorily significant features of the world. Roughly speaking this is the task of arranging our language so that our linguistic categories ‘cut the world at its joints’.¹²

If we view scientific progress as involving the accommodation of language to the causal structure of the world, then we need a theory of reference that provides a plausible account of capturing the causal structure of the world. Providing such an account also involves providing an account of failing to ‘cut the world at its joints’. We now believe that at least some of the terms of discarded theories do not refer—for example, ‘phlogiston’, ‘optical aethers’, ‘caloric’, and ‘demon’. An adequate realist theory of reference must account for both reference success and reference failure.

Nonetheless, reference failure has not received sufficient attention from causal theorists of reference. This neglect, however, is not surprising. One of the main motivations behind causal theories of reference is to explain how the *terms* of science can refer to the same aspects of reality when the *theories* about those aspects of reality change radically. One of the main objections to description theories of reference is that they have the unwelcome consequence that revolutionary theory change results in a change in the referential properties of theoretical terms. If the reference of a term is determined by a cluster of descriptions associated with the term, and if the cluster of descriptions changes significantly as the result of a scientific revolution, then it

Fine assumes that Laudan’s objections [1981] to the simple abduction are conclusive. Indeed, his main objection to Boyd relies on this assumption (Fine [1984], 84, 88, 89). The arguments that follow are thus relevant to Boyd’s methodological argument for scientific realism. For other objections to abductive arguments for realism see Bas van Fraassen *The Scientific Image* (Clarendon Press, Oxford, 1980). For an excellent and insightful recent discussion of the whole debate see Alison Wylie ‘Arguments for Scientific Realism: The Ascending Spiral’ (*American Philosophical Quarterly*, 23, [1986], pp. 287–97).

¹² Boyd ([1979], p. 358).

seems to follow that the term will refer to something different in the new theory. But, this implies that the two theories are about different subject matters, in which case the two theories do not really conflict at all. It is perhaps Kuhn's most famous claim that revolutionary theory change results in a change of subject matter, and thus the theory that precedes and the theory that succeeds a scientific revolution are incommensurable and rationally incomparable.¹³ To avoid this consequence we need a theory of reference that allows a successor theory to be about the same aspect of reality as its predecessor; this in turn requires a theory of reference that allows us to be largely wrong about the entities to which we refer. Defenders of causal theories of reference view this feature of their approach as one of its central virtues.

As a result of this emphasis, however, causal theorists have not dealt sufficiently with the issue of reference failure. Many central terms of abandoned theories are now thought to be referentially empty. How can causal theories of reference make sense of reference failure for such theoretical terms? That is, how can we distinguish cases where we successfully refer but are largely (or completely) wrong about the entity referred to from cases of reference failure? Since causal theories of reference allow us to be largely (or completely) wrong about what we refer to, it is far from clear how such theories can allow for reference failure. To demonstrate this claim, and to exhibit the lack of attention this problem has received, let us look at the account of reference failure provided in the most involved attempt to work out the *details* of a causal theory of reference.¹⁴

Michael Devitt claims that, according to the causal theory of reference, a theoretical term refers 'to those aspects of reality that prompted its use'.¹⁵ How on this account can a theoretical term fail to refer? Though Devitt does not want 'phlogiston' to refer to oxygen, according to our best science it was oxygen that caused people to come up with phlogiston theory; hence oxygen prompted us to come up with the term 'phlogiston'. In *Designation*, Devitt writes that in some cases 'there may not be *any* aspect of reality that prompted the use of the term.'¹⁶ In such cases the term is empty; it has no referent. But clearly, in all relevant cases, *some* aspect of reality prompts the use of the term. If one, uncharitably, takes Devitt's claim literally then, such terms as 'phlogiston' or 'optical aether' would not be empty after all. Devitt does recognize a problem here; his solution, however, is inadequate. In conversation he has suggested that he would amend the above statement to say 'there may not be *any appropriate* aspect of reality that prompted the use of the term'.

¹³ See Kuhn *The Structure of Scientific Revolutions* (Chicago 1962, second edition 1970), and *Criticism and the Growth of Knowledge* Lakatos and Musgrave (ed.) (Cambridge, 1970).

¹⁴ For those theorist most instrumental in developing the causal approach see footnote 2. Michael Devitt *Designation* (Columbia, 1981), however, has provided the most complete fleshing out of the details of a causal theory of reference.

¹⁵ Devitt 'Science and Truth,' unpublished notes ([1982], III, 3).

¹⁶ Devitt [1981], p. 201.

But this will not do. Without an explanation of what counts as an ‘appropriate’ aspect of reality, we can not determine the informativeness of the claim that ‘terms of a mature science typically refer’. Indeed, if the account of reference fixing is *restricted* to what, in reality, *prompted* the use of a term, then theoretical terms cannot fail to refer. As a consequence, the referential claim of scientific realism would become trivial. The claim that some aspect of reality, we know not what, prompts the use of theoretical terms is not worth fighting for. The causal theorists of reference must supply criteria of appropriateness, otherwise the referential claim of scientific realism is uninteresting. More specifically, we need some indication of how reference fixing is tied to success in the endeavor of ‘cutting nature at its joints’.

Devitt’s main goal is to develop an account of reference fixing which avoids the Kuhnian conclusion that major changes in scientific theory do not involve greater insight into the nature of the world but are instead changes of subject matter. In Section 3 I develop a theory of reference for theoretical terms which both allows our theories to be wrong, in important respects, about the entities they refer to and which also allows for the possibility of reference failure. I argue that, if we are to give an adequate account of reference for theoretical terms, we must consider not only what prompted the use of a term but also what the theorists say in using the term. The descriptions must play more of a role than causal theorists have thus far allowed. Theory constitutive metaphors provide the descriptive constraint limiting the referent of the term to an ‘appropriate’ aspect of reality.¹⁷

3 REFERENCE FAILURE AND THEORY CONSTITUTIVE METAPHOR

The causal theory of reference, with its emphasis on the compatibility of massive error and successful reference, is most plausible for terms referring to ostensible objects whose real essences consist of features of internal constitution. Many scientific terms, however, are completely theory embedded. For various reasons, many theoretical terms refer to entities which are unobservable or, to use Enc’s alternative, non-ostensible.¹⁸ As McMullin has incisively put it, ‘electrons are what quantum theory says they are.’¹⁹ The only reason we have for believing that there are electrons is the success of quantum theory.

¹⁷ Barent Enc’s ‘Reference of Theoretical Terms (in *Nous* 10, [1976], pp. 261–82) also defends the conclusion that a pure causal theory of reference cannot account for the explanatory function of new theoretical terms. His account, however, does not incorporate or respond to the concerns of causal theorists. Since Enc argues that the new terms must be associated with a set of sentences which specify the ‘identity conditions’ of a new kind (p. 280), his account appears to be less flexible than the metaphor approach. On the other hand, his emphasis on an ‘explanatory mechanism’ associated with the newly named kind (pp. 271, 277) fits in nicely with the metaphor approach. As we shall now see, the metaphor approach emphasizes the open ended, tentative, and programmatic nature of scientific research.

¹⁸ Enc ([1976], p. 269).

¹⁹ McMullin ([1984], p. 21).

Indeed, some of the most important theoretical terms refer to kinds whose 'real essences consist of complex relational properties' rather than features of internal constitution.²⁰ Boyd's discussion of 'theory constitutive metaphors' provides a plausible account of the descriptive element associated with these theoretical terms.

The basic idea is that new theoretical terminology is introduced in conjunction with explanatory metaphors. A metaphor involves a primary subject which is well understood and a secondary subject which is thought to have interesting similarities to the primary subject. (Note: 'primary' refers to primacy in understanding rather than the subject of primary interest to us; that which we are interested in understanding is the 'secondary' subject of the metaphor.) The specific nature of these similarities, however, displays what Boyd calls 'inductive open-endedness' and what Ernan McMullin calls 'tentative suggestion'. Boyd writes,

The reader is invited to explore the similarities and analogies between features of the primary and secondary subject, including features not yet discovered or not yet fully understood . . . Theory constitutive metaphors are introduced when there is (or seems to be) good reason to believe that there are theoretically important respects of similarity or analogy between the literal subjects of the metaphors and their secondary subjects. The function of such metaphors is to put us on the track of those respects of similarity or analogy.²¹

In short, theory constitutive metaphors are 'invitations to future research': they serve to introduce terms without specifying the defining characteristics of the referent of the term, but instead leave it open for the routine business of scientists to discover the yet unspecified essential properties of the referent in question.

The introduction of theory constitutive metaphors is an essential part of the initial baptism or dubbing ceremony of many theoretical terms. As a consequence, the introduction of the theoretical term, though inductively open-ended, does involve 'some tentative or preliminary indication of the properties of the presumed kind in question'. More specifically, there is the assumption that the referent of the term will prove to be similar to the primary subject, and will thus vindicate the claim to reference. As a corollary,

In cases where a theoretical metaphor proves not to represent a real insight, we need no more inquire about the new referents of its metaphorical terms than we do with respect to the referent of 'vital force': in such cases the 'guess' does not work out, and the relevant terms do not refer at all.²²

This account of reference failure follows quite naturally from the 'programma-

²⁰ Boyd ([1979], p. 358).

²¹ Boyd ([1979], p. 363) and McMullin ([1984], p. 31).

²² Boyd ([1979], p. 370).

tic inductive open-endedness' of theory constitutive metaphors. Confronted with something like human cognition, we introduce the term 'information processing' assuming a theoretically interesting analogy with machine computation. Whether we have indeed succeeded in referring to some aspect of reality will depend on whether or not the new research program proves to be fruitful—*i.e.*, it depends on whether subsequent science discovers that human cognition is analogous in theoretically important respects to machine computation.

The explanatory metaphor, which is associated with the new theoretical term, restricts the scope of future acceptable error. A scientific term like 'optical aether' is supposed to provide a linguistic category which captures and refers to causally and explanatorily important features of the world. Scientists introduce new terminology in order to more adequately cut the world at its joints. The idea is not just to introduce a term which will refer to whatever is responsible for certain empirical phenomena. The term 'optical aether' is introduced and used by Fresnel, MacCullagh, and Maxwell with more specific intentions. Optical aethers were supposed to provide a medium of transmission of transversal optical waves. The manner in which this medium transmitted optical waves was supposed to be similar to waves in a liquid. If optical aethers were similar to a liquid, then light could be explained by analogy to waves in a liquid. This would provide a purely mechanical account of light on a par with the mechanical rising and falling that constitutes the crest and trough of a wave propagated through a medium. The term is essentially linked to this metaphor, because if it were not, scientists would not achieve a greater accommodation of language to the causal structure of the world.

It might be suggested that the idea of categorical sortals, rather than metaphors, provides the necessary descriptive element. In the case of optical aether, one might argue that the sortal involves the concept of a 'medium of transmission'. This approach is attractive. I have come to believe, however, that it is too close to the pure causal theory's 'whatever causes the phenomenon in question' approach. When a scientist introduces a new theoretical term, the point is to name a new explanatorily significant aspect of the world. As Enc has argued, in cases involving non-ostensible entities, 'the only reason we have for supposing that the entity exists is that we believe that our supposition will help us explain a set of phenomena. And these explanations can proceed only by our supposing that the entity has a set of properties and by our hypothesizing an explanatory mechanism which would show how the entity, given the properties it is supposed to have, brings about its effects.'²³ The idea that light involves some medium of transmission, I know not what, does not pack enough explanatory punch. The explanatory mechanism must also play a role in reference fixing.

²³ Enc ([1976], pp. 276–7).

The case of phlogiston and oxygen quite clearly reflects the inadequacy of the sortal approach. The sortal associated with 'phlogiston' is the category of a new substance. Oxygen, however, is a new substance and yet phlogiston is the paradigm of reference failure. If one is committed to the sortal approach, then one will shrug and assert that 'phlogiston' referred to oxygen. But why bite this bullet? The reason scientists came to believe in oxygen and not phlogiston is that the explanatory mechanism, the metaphor of a substance emitted from burning objects, was rejected and the new explanatory mechanism associated with oxygen was accepted. Categorical sortals without the hypothesized explanatory mechanism do not fill the bill. Typically, of course, the explanatory mechanism will presuppose a categorical sortal, or general kind term, and the discovery that there is nothing of that kind causing the phenomena will undermine the suggested explanatory mechanism. Sortal failure typically entails mechanism failure. Mechanism failure, however, does not entail sortal failure.

Nonetheless, one may still appeal to a categorical sortal *and* an explanatory mechanism and thereby avoid the messy notion of metaphor. So why, one might ask, identify the explanatory mechanism with a theory constitutive metaphor? 'Optical aether' was introduced with the intention to refer to something like a liquid. Thus, one might say we need a more specific kind of sortal which links the term more closely to the explanatory mechanism; for example, 'optical aether' refers to a liquid medium of transmission.

The category liquid, however, is an even more inadequate sortal because 'optical aether' is not supposed to literally refer to a liquid. Indeed, optical aether cannot be just like a liquid, for it must be impervious to all causes except propagations of this one kind of energy through it. Optical aether is supposed to be like a liquid but not a liquid. Furthermore, despite the assumption that optical aether will turn out to be like a liquid, scientists could not explain the ways in which this new medium of transmission will be like a liquid and the ways in which it will be different. The essential properties of this new substance were yet to be discovered. It is the task of the ensuing research program to discover the essential properties of the like-a-liquid-but-not-a-liquid medium. To use Boyd's terminology, the optical aether explanatory hypothesis was 'inductively open-ended'. Nonetheless, since the intention was to refer to something like a liquid, the term could successfully refer, only if the relevant phenomenon is indeed a medium analogous to the primary subject. In the case of the term 'optical aether' the suggestion did not prove fruitful. According to our current best science, this aspect of the world is not relevantly similar to the primary subject of the metaphor.

The concept of metaphor is thus to capture the elements of analogy, mystery, and open-ended or tentative suggestion involved in the introduction of new theoretical terminology. The relationship between the success of theories and the reference of theoretical terms will be discussed below. First,

however, we need to recognize the compatibility of this theory with the central (scientific) realist insights of the causal theory.

Causal theories of reference introduced the idea of initial baptism or dubbing ceremonies, in which terms were grounded to objects, to take the place of definitional accounts of reference fixing. It may seem that by moving away from 'dubbing' or 'baptism' one has lost the link between language and the world that it provided. Kuhn, for example, raises this objection in response to Boyd. Kuhn argues,

However imperfectly developed, 'dubbing' was introduced in an attempt to understand how in the absence of definitions, the referents of individual terms could be established at all. When dubbing is abandoned or shoved aside, the link it provided to the world disappears as well.²⁴

Though this line of objection is well taken, I believe that it can be answered.

We were forced to introduce the above account of reference fixing via theory constitutive metaphors to avoid the alternative view that a theoretical term refers to *whatever* aspect of reality prompted its use. More specifically, we needed to explain how, in cases of reference failure, no 'appropriate' aspect of reality prompted the use of the term, and thus the term was empty. Kuhn questions how this approach can ground the names of natural kinds. In the case of many natural kind terms, we are able to point to a sample and thereby dub all things with the same essential internal constitution. In the case of the theoretical terms we have been discussing, however, the problem is more acute. How can we ground the term 'electric charge' to a physical magnitude, whose real essence involves a complex set of relational properties, by baptizing a single (or single type of) manifestation of this physical magnitude? The solution defended in this paper appeals to 'theory constitutive metaphors', but Kuhn's fear is that this solution abandons the link to the world that 'baptism' or 'dubbing' ceremonies provide. But why must we abandon this link with the world?

Theory constitutive metaphors are introduced so that theoretical terms will refer to the cause of some phenomenon confronting scientists. As a result it is linked to the world through the scientist's confrontation with the effects of the entity doing the causing. There is no reason to abandon the *causal* link to the world. The point of the employment of metaphor is to add the *additional* specification (though tentative, programmatic, inductively open-ended) necessary to ground the term to a particular causally interesting aspect of reality. Furthermore, theory constitutive metaphors enable us to lock the term

²⁴ Kuhn 'Metaphor in Science' in *Metaphor and Thought*, Andre Ortony (ed.) (Cambridge University Press, 1979); p. 412. Kuhn focuses his concern on Boyd's use of 'epistemic access'. In order to avoid interpreting and discussing this notion (a worthwhile paper in its own right), I have left Boyd's use of epistemic access' out of my discussion.

onto a natural family and not just one (or some) of the individuals that constitute the natural family.

Devitt has suggested that theoretical term are grounded to the world through a form of 'quasi-perceptions' in which we are 'perceiving the referent through the instrument'.²⁵ The idea is that when we lack direct ostensive access to the entity or kind to which we are trying to refer, we view the term as grounded via our instruments to some aspect of reality. Similarly, Hacking and McMullin defend scientific realism against problems of Kuhnian meaning change by emphasizing the causal link from researcher to entity via instrument.²⁶ We shall view 'quasi-perception' through instruments as still a necessary condition of the reference grounding process. However, because of the theoretical nature of the referent of the term, the additional specification provided by the theory constitutive metaphor is also necessary to link the term to an appropriate aspect of reality.²⁷ Finally, if the metaphor proves to be fruitful, then we are justified in believing that we have successfully referred to a causally significant feature of the world.

The metaphor plus causal interaction analysis, of course, does not require that all speakers who use a theoretical term know of the metaphor associated with the term. As Dummett and Putnam have emphasized, there is an important division of linguistic labor.²⁸ In fixing the reference of many theoretical terms a metaphor is associated with the term which gives a tentative and preliminary indication of the properties of the kind in question. This information need not, however, be possessed by all those who are able to successfully use the term. It is only necessary that those who use the term are part of a linguistic community which is composed of some individuals who are experts about the theoretical term in question. Within such a linguistic community, there is a sense associated with the term which plays an important role in determining the reference of the term, but one need not know this sense to successfully use the term.

Thus, *if* one maintains that a 'description' theory of reference requires that in order to refer successfully one must know a description, or cluster of descriptions, which is sufficient to uniquely pick out the referent of the term, then there is an important contrast between what is required by the theory

²⁵ Devitt ([1981], p. 201).

²⁶ Hacking ([1984], pp. 159–61) and McMullin ([1984], pp. 22–3).

²⁷ Boyd argues that his concept of 'epistemic access avoids the necessity for idealized dubbing ceremonies. The kind to which a general term refers is determined by the role that the term plays in socially coordinated inquiry' ([1979], p. 388). My arguments are compatible with Boyd's extremely plausible conclusion. Quasi-perception plays an important causal role in the socially coordinated activity.

²⁸ For a more complete discussion of this division of labor see Boyd ([1979], pp. 388–91), Putnam ([1977]), pp. 227–9), and Dummett ([1973], pp. 137–41). Despite the association between the division of linguistic labor and Putnam's causal theory of reference, it can be utilized by sense theories of reference as well. Indeed Dummett [1973] has argued that Frege incorporated just such a principle in his theory of reference.

here defended and description theories. Nonetheless, this account of reference does claim that there are descriptions associated with many theoretical terms that place limits on what can count as the referent of the term; there are certain ways in which we (the linguistic community using the term) cannot be wrong about the properties of entities to which we allegedly refer.

Some have thought that the major argument for the causal theory of names is that it allows for the possibility that one can be mistaken in *all* important respects about the entities to which we refer. When we conclude that nothing has the properties attributed to Jonah it allows for the possibility that Jonah existed and that we have been all wrong about him (her or it).²⁹ Description theories of names do not countenance this possibility, but instead compel us to conclude that 'Jonah' does not refer. The above account of reference for theoretical terms, similarly, does not allow for such a possibility. In this respect there is a clear difference between the metaphor approach and the pure causal approach. To many, however, this will be a welcome conclusion. As Boyd has pointed out, the main reason for sustained interest in Law Cluster accounts of theoretical terms is the judgment that 'there is something absurd in the suggestion that all of our most fundamental beliefs about a "theoretical entity" might be fundamentally mistaken.'³⁰ Whether one is attracted to the metaphor approach or the sortal approach or some other approach, there simply must be limits on the scope of possible error when it comes to theoretical terms. This is a necessary result for any theory which accommodates the scientific belief that 'phlogiston', 'aether', 'caloric', 'demon', etc. did not refer, because there was no *appropriate* aspect of reality that prompted their use.

It might be objected that there is no determinate answer to the question whether a theoretical term (i) referred and the theorist were largely wrong about its essential properties or (ii) did not refer at all. The metaphor approach clearly leaves open the possibility of indeterminate cases. Nonetheless, if the standard cases of reference failure are clear and determinate enough, the possibility of indeterminacy is not problematic. Indeed, if the metaphor approach provides an account of reference fixing which accounts for the widespread agreement among scientists over reference successes and failures, then the possibility of indeterminate cases does not tell against the account. We should expect that it will sometimes be difficult to tell whether an explanatory metaphor has been fruitfully articulated or been abandoned for a new explanatory mechanism.

This issue, however, is complicated by the successes of referentially empty theories. As Larry Laudan has emphasized, many theoretical terms that we now believe are referentially empty were embedded in theories that were quite successful. The optical aether theories discussed above, for example, enjoyed

²⁹ See Kripke ([1980], p. 66–7); Devitt ([1981], pp. 13–20).

³⁰ Boyd ([1979], p. 387).

considerable success.³¹ If there is no optical aether, how are we to account for the success of the theory? Indeed, if theories with referentially empty terms can be successful, how can we ever justifiably believe that our theoretical terms refer to a causally significant feature of the world? To answer this question, we must distinguish success which provides evidence for reference and success that provides no evidence for reference. The success of aether theories, for example, in no way clarified the mystery of the nature of aether. As Hardin and Rosenberg have argued, aether theories demonstrated only that 'one could give a consistent mechanical model for the bearer of light vibrations.'³² A mechanical model can be successful despite the utter mystery of the actual mechanism. We thus need to distinguish experimental successes in accounting for empirical phenomena (let us call this 'phenomenal success') and success in discovering the essential nature of the new kind (let us call this 'articulation'). Let us consider another familiar example.

Caloric theory is a good example of a successful theory with a central term which failed to refer. In order to explain the phenomenon of heat, scientists introduced 'caloric' as a term they understood to refer to some kind of subtle, conserved, fluid-like stuff. The basic idea was that caloric, this conserved stuff, was held in other stuffs in a manner similar to water in a sponge. The term was grounded to the world via the scientists' ability to measure heat, and otherwise experience heat, and the term allegedly referred to something they took to be analogous to a subtle fluid.

Caloric theory was quite successful in heat transfer experiments. Indeed, caloric theory generated a substantial research program and body of theory dealing with melting, boiling, and thermal expansion. Eventually, however, the interconvertability of heat and work showed that heat was not a conserved stuff. The non-conservation of heat undermined the caloric theory and paved the way for the kinetic theory of heat. Despite the phenomenal success of the caloric theory, the supposition that there was a fluid-like stuff, which was responsible for heat, eventually became untenable. The new (or return to the old) theory that heat was molecular motion or energy accounted for the non-conservation of heat and, since energy is conserved, was equally suitable for explaining the complex array of heat transfer experiments. In addition to being extremely fruitful as a research program, the supposition that heat was molecular motion cohered nicely with the developing molecular theory.

The phenomenal success of the research program based on the supposition that heat is energy has also been accompanied by a further articulation of the nature of heat and of how heat transfer occurs. In contrast, the supposition that heat was a subtle fluid, despite the phenomenal successes of the research program, was not successfully further articulated and thus, in this respect, the metaphor did not prove to be fruitful. The caloric theory, unlike the kinetic

³¹ Laudan ([1981], p. 27).

³² Hardin and Rosenberg ([1982], p. 612).

theory, did not lead to a deeper understanding of the essential properties of the kind in question.

Both caloric and optical aether theories are excellent examples of success which provides strong evidence of reference (articulation) and success which warrants a more instrumentalist interpretation of a theory (mere phenomenal success). First, the development of both theories provided a useful foundation for the emergence of their successor theory. The results of heat transfer experiments were compatible with both the caloric theory and the kinetic theory of heat. So, the non-conservation of heat rendered the kinetic theory the most plausible successor theory. The success of optical aether theories supported the adequacy of a mechanical model for the bearer of light. But the success of the aether theory provided equal evidence for a quantum theory of light. In each case, the theories made referential claims involving explanatory mechanism which future research did not support. The ontological shift between the theories does not undermine the claim that the development of the theories reveals a zeroing in on nature's joints.

Second, theory constitutive metaphors introduce a research program where the scientist is to discover and attempt to further articulate the theoretically important respects of similarity and analogy between the primary subject and the secondary subject. It is the success in discovering and articulating the similarities between the primary subject and the secondary subject which warrants a realistic interpretation of the theories' ontological commitments. As in the cases of caloric and aether theories, there may be phenomenal success (which clearly implies that you have something right) without having success in the project of further articulating the essential nature of caloric or optical aether. There is a clear and significant difference between the scientists' ability to articulate the essential nature of the explanatory mechanism of the kinetic theory and their ability to articulate the essential properties of the substance caloric. Similarly, the mysterious, and seemingly impossible, aether medium was never coherently articulated. The inability of a research program, despite its phenomenal successes, to further articulate its explanatory, theory constitutive, metaphor undermines the referential claim of the central term. When a new explanatory mechanism can explain the successes of the old theory and can be more fully articulated, it is reasonably concluded that no explanatorily significant feature of the world was picked out by the central term of the old theory.

The phenomenal success clearly counts in a theory's favor but the less we understand about the explanatory mechanism suggested by the metaphor, the less justified we are in believing that the central term refers. If the rough, tentative, programmatic, inductively open-ended specification is further *developed and articulated* by the subsequent history of a *successful* theory using the term in question, then we are surely justified in believing that the term has indeed locked onto a causally and explanatorily significant feature of the

world. On the other hand, the more difficult it is to discover the essential nature of the kind in question, the less justification we have for our belief in its existence.

Of course, when theories are flourishing it is not wrong-headed to have practical confidence in them despite their many mysteries. Indeed, it makes sense to stick with a phenomenally successful theory because success implies that one has got something right. Confidence in a theory, however, is compatible with awareness of its inarticulated, and perhaps, mysterious aspects. It is, of course, part of the ongoing project of science to work out the tentative and inductively open-ended suggestions of the theory. The epistemic warrant for an ontology corresponds to the success of this aspect of the research program. This suggests that even when a theory is in its heyday a degree of agnosticism about the realism of its ontology may be the correct epistemic frame of mind. If a research program consistently fails to articulate the essential nature of its purported ontology, then an instrumentalist attitude is called for. But, of course, if it is the best phenomenal game in town, it clearly makes practical sense to stick with the theory.

Larry Laudan has provided an impressive list of theories which we would now claim are non-referring, yet which he claims were nonetheless successful theories.³³ I believe that most (and perhaps all) of the theories on the list would prove, under close scrutiny, to be similar to caloric theory and optical aether theory (two of Laudan's candidates) in providing support for the theory of reference here defended and support for the view that the development of science reveals an accommodation of linguistic categories to the causal structure of the world.

Since I have not discussed how in each of Laudan's cases the attempts to further articulate the metaphors proved to be unfruitful, the historical aspect of the argument is incomplete. Nonetheless, many of the historical details have already been presented by others. Hardin and Rosenberg argue that the crystalline sphere theory of astronomy failed to provide a comprehensive and coherent model for all planetary motion.³⁴ Such an 'anaemic' astrophysics is no cause for realist concern. McMullin argues that modern geology, cell biology, and chemistry exhibit a progressive discovery and articulation of the

³³ Laudan offers the following already infamous list:

- the crystalline spheres of ancient and medieval astronomy;
- the humoral theory of medicine;
- the effluvial theory of static electricity;
- 'catastrophist' geology, with its commitment to a universal (Noachian) deluge;
- the phlogiston theory of chemistry;
- the caloric theory of heat
- the vital force theories of physiology;
- the contact-action gravitational aether of Fatio and LeSage;
- the electromagnetic aether
- the optical aether (Laudan [1981], p. 33).

³⁴ Hardin and Rosenberg ([1982], p. 610).

hidden structure of the world.³⁵ The careful and clear details of the many theories in these fields contrast strikingly with the mysteries of their predecessors: catastrophist geology and the universal Noachian deluge; vital force theories; and the phlogiston theory of chemistry. Similarly, humoral theories of medicine did not provide a clear or comprehensive explanatory account of the domain it claimed. It is thus fair to conclude that the successes of referentially empty theories do not tell against the realist thesis that science is converging on an accurate accounting of the causally significant features of the world.

Finally, in addition to cases of reference failure and successful reference, the metaphor approach allows for cases of partial reference. Indeed, Boyd incorporates Hartry Field's semantics of partial reference into his theory of reference.³⁶ In many cases we will succeed in referring only to some aspect of reality that subsequent research will show to be complex in nature, and thus further accommodation of language will be necessary if we wish to 'cut the world at its joints'. The paradigm case of partial reference, as developed by Field, is Newtonian mass and post-Newtonian relativistic mass and proper mass. Field argues that the Newtonian term 'mass' partially referred to two physical quantities: 'relativistic mass' ($= \text{total energy}/c^2$) and 'proper mass' ($= \text{nonkinetic energy}/c^2$).

Another example is the natural kind term 'acid'. It turns out, according to our current best science, that what Lavoisier, Dalton, and Davy called 'acids' and theorized about (often falsely) was not one but two natural kinds: Bronstead-Lowry acids and Lewis acids.³⁷ The problem is not one of common ambiguity. There simply is no matter of fact which determines which of the two natural kinds Dalton, for example, was referring to when he used the term 'acid'. Field shows that the only plausible solution is that Dalton partially referred to each of the two natural kinds.³⁸ In some cases, we directly dub various examples of the alleged natural kind and then later discover the need

³⁵ McMullin ([1984], pp. 27–9).

³⁶ See Boyd ([1979], p. 378ff.). On the semantics of partial reference see Hartry Field 'Theory Change and the Indeterminacy of Reference' (*Journal of Philosophy*, 70 [1973], pp. 462–81). I assume that Field's arguments are correct.

³⁷ For the details see Hacking ([1983], pp. 84–5).

³⁸ Again see Field ([1973]: esp. pp. 467–74). Field provides the following Tarski style semantics for partial reference: 'In order to give a semantics for indeterminate expressions, let's introduce the term "structure"; a *structure* for a sentence is a function that maps each name or quantity, and maps each predicate into some set. The structure *m* corresponds to the sentence if each name or quantity term of the sentence partially denotes the thing that *m* assigns to it. Now, for each structure *m*, we can apply the standard referential (Tarski-type) semantics to determine whether the sentence is *m-true* or *m-false*, i.e. true or false relative to *m*. (To say that the sentence is *m-true* is to say it would be true if the denotations and extensions of its terms were as specified by *m*.) We can then say that a sentence is *true* (false) if it is *m-true* (*m-false*) for every structure *m* that corresponds to it. Putting all these definitions together, we get definitions of truth and falsity in terms of partial denotation and partial signification' (Field [1973], p. 477).

for additional accommodation of language to the causal structure of the world. The same holds true for terms which are introduced with theory constitutive metaphors. Cases of partial reference and bifurcating natural kinds do not undermine scientific realism but instead lend support to the thesis that the development of science reveals an accommodation of language to the causal structure of the world.

The metaphor approach is a non-definitional account of reference fixing which incorporates the strengths of both the pure causal approach and pure description theories. This hybrid approach justifies the realist position that the development of a scientific theory involves the discovery of essential properties rather than changes of subject matter. As long as an explanatory metaphor proves to be insightful, then the changes in our views about the essential properties of the referent can be viewed as greater insights into the causal structure of the world. We may say, *contra* Kuhn, that a term is coreferential across paradigms even if the change in paradigm dictates the need to say quite different things about the referent of the term. Furthermore, we can view the world as dictating the changes in our views about the properties of the referent. In the terms employed by Devitt, it is that aspect of reality that initially prompted the use of the term that is responsible for the subsequent changes in the properties we attribute to the referent of the term. Thus, although this account of the grounding ceremony does involve a modification of the paradigm case of grounding via perception (developed by Devitt), it nonetheless seems to capture the major motivational impetus of the causal approach to reference.

As we saw in Section 2, causal theories of reference fail to provide an account of reference failure for theoretical terms and thereby trivialize the referential claim of scientific realism. If a causal theory of reference is to avoid this consequence, then the descriptions associated with some theoretical terms must play a role in fixing the reference of the term. Whatever the shortcomings of the metaphor account of reference fixing, it appears that only a hybrid theory of reference will allow for both reference failure and improved insight; only such a theory will suit the needs of the scientific realist.

4 SCIENTIFIC REALISM AND THE META-INDUCTION

With the metaphor approach in hand let us return to the claim that 'terms of a mature science typically refer' and evaluate the adequacy of the meta-induction. The meta-induction is quite simple:³⁹

³⁹ This reconstruction of the meta-induction is from a draft of Devitt's *Realism and Truth* (Princeton, 1984). Devitt's final version is as follows: 'At any time *t* in the not too immediate past it would have been a mistake to infer Scientific Realism (*t*) from what science at *t* (apparently) posited to explain observable phenomena, for it turned out not Scientific Realism (*t*). So it is probably a mistake to infer Scientific Realism (now) from what science now (apparently) posits to explain observed phenomena' ([1984], p. 145).

1. For any past theory, *t*, most of the theoretical entities posited by *t*-science do not exist (according to our current best science).
2. So, probably (according to some future science), most of the unobservables posited by our current best science do not exist.
3. Scientific realism asserts that most of these entities do exist.
4. So, probably, scientific realism is false.

As it stands this argument is not telling against the scientific realist. As I have been emphasizing, what matters for the scientific realist is the increasingly 'accurate' accommodation of language to the causal structure of the world'. Since the realist does emphasize the convergent progress of science, the realist does wish to claim that in a *mature* science the terms typically refer. The metaphor approach provides a straightforward explication of 'mature' that I believe undermines the sceptical power of the meta-induction.

In explicating the claim that the terms of a mature science refer our concern is with the level of articulation of *particular theories* which make up the science in question. According to the metaphor approach, we are justified in believing that a theoretical term successfully refers when four necessary conditions are met: (1) quasi-perception provides a causal link to the world; (2) a theory constitutive metaphor provides a link to an underlying trait property or natural kind/family; (3) the resulting research program is phenomenally successful; and (4) the explanatory metaphor is fruitfully articulated (this last condition is met when the research program, based on the inductively opened initial specification, is successful in further articulating the similarities between the literal subject and the secondary subject and thereby the essential nature of new explanatory kind is more fully discovered). With this account of reference in hand, we can disarm the meta-induction.

The scientific realist need not maintain that most of the terms of contemporary science typically refer. For realist purposes it is sufficient to maintain that the central terms of mature theories typically refer. When 'mature' is taken to qualify particular theories, theories whose central terms meet the above conditions upon reference, then the plausibility of the meta-induction is significantly diminished. Indeed, if one further clarifies the realist position by also including cases of partial reference, it is hard to believe that the meta-induction can be sustained.⁴⁰

Finally, the argument of this paper supplements, in an important way, Devitt's response to the meta-induction in *Realism and Truth*. Devitt suggests that the meta-induction can be undermined by suspending judgments on *the frontier of science* (at any given time) and also in cases where there is dispute within the scientific community. Thus, the issue becomes how many entities

⁴⁰ On the power of partial reference as a tool for blocking the meta-induction see Devitt 'Against Incommensurability' (*Australasian Journal of Philosophy*, 57 ([1979], pp. 29–50); Devitt ([1984], pp. 147–49, 156; and of course Field [1973]).

past science *confidently posited* and similarly with current science.⁴¹ Although my analysis is sympathetic to this general line of response, I believe that the account in this paper provides a more clearcut and objective criterion of the terms or referents we wish to count when deciding whether or not science has been zeroing in on nature's joints. Scientists may 'confidently posit' entities when they cannot satisfy the conditions necessary for reference, and there may be entities posited by theories on the 'frontier' of science that are part of what I have characterized as mature theories. Most importantly, the theory of reference has been connected to the work of scientists striving to grasp the causal structure of the world; as a consequence, we get a realist response to the meta-induction motivated by and connected to a plausible explanation of how working scientists discover the unobservable 'joints of nature'.

In conclusion, we should expect (i) reference failures, like that of caloric theory; (ii) situations of partial reference, like that of Newtonian mass and post-Newtonian proper mass and relativistic mass; and (iii) situations, like Bohr's model of the atom, which reveal a long process of refinement and further articulation of a model or metaphor.⁴² These sorts of cases are to be expected in the slow process of zeroing in on nature. As long as cases of reference failure do not directly indicate that we are not cutting ever more accurately at nature's joints, they should not cause alarm. Cases of reference failure should not induce skepticism about the value of science as a machine for uncovering nature's hidden ontology. As science uncovers more of the causal structure of the world, it is natural to expect a continuing evolution of the language of science. The lesson to be learned from the meta-induction is that the process of scientific discovery and linguistic evolution is still going on.

*Department of Philosophy and Religion
Bates College
Lewiston
Maine 04240*

⁴¹ See Devitt ([1984], pp. 17, 122, 144).

⁴² Bohr's model of the atom is discussed throughout Boyd's article; to avoid repetition I have concentrated on cases which are not extensively discussed by Boyd.