The Semantics and Metaphysics of Natural Kinds

Routledge Studies in Metaphysics

1. The Semantics and Metaphysics of Natural Kinds
Edited by Helen Beebee and Nigel Sabbarton-Leary

The Semantics and Metaphysics of Natural Kinds

Edited by Helen Beebee and Nigel Sabbarton-Leary

Routledge
Taylor & Francis Group
New York London

First published 2010
by Routledge
270 Madison Ave, New York, NY 10016

Simultaneously published in the UK
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Typeset in Sabon by IBT Global.
Printed and bound in the United States of America on acid-free paper by IBT Global.

Library of Congress Cataloging in Publication Data
A catalog record has been requested for this book.

ISBN10: 0-415-87366-5 (hbk)
ISBN10: 0-203-85233-8 (ebk)

ISBN13: 978-0-415-87366-6 (hbk)
ISBN13: 978-0-203-85233-0 (ebk)

Contents

Acknowledgments

The editors would like to thank the Arts and Humanities Research Council, whose financial support (for the project *Metaphysics of Science*, AH/D503833/1) made the production of this volume possible. We would also like to thank the Leverhulme Trust, whose financial support enabled Helen Beebee to complete her contribution.

We would also like to thank the contributors for making our lives much easier than they might have done. Finally, thanks to Mia Sabbarton-Leary and Gavin Brown, for all kinds of reasons.

1 Introduction

Helen Beebee and Nigel Sabbarton-Leary

The topic of natural kinds is one that has been much discussed in metaphysics, philosophy of science, and philosophy of language in recent decades. This collection brings together contemporary work in these areas, in the hope that doing so will highlight how views and issues in one area affect those in other areas. For example, some philosophers of language hold that a semantic theory of 'natural kind terms' should aim to capture a class of general terms that designate natural kinds in the metaphysician's sense—excluding terms such 'bachelor' and 'pencil' but including terms such as 'gold' and 'water'. Metaphysicians, in turn, often appeal to semantics—and in particular to the Kripke-Putnam account of natural kind terms—in order to carve out a metaphysically substantial class of necessary truths. And some philosophers of science take a Kripke-Putnam-style causal theory of reference to defeat Kuhnian relativism.

1. CLASSIFYING NATURE: THE METAPHYSICS PERSPECTIVE

One source of philosophers' interest in natural kinds is distinctively metaphysical: are there 'natural joints' in nature, which our classificatory systems—in ordinary life or in science—might latch onto? For example, compare general terms such as 'cat', 'silver', carbon', 'electron', and 'planet', on the one hand, with expressions such as 'object bigger than a car', 'creature with four or more legs', and 'carbonated drink' on the other. The members of the former list of terms, but not the members of the latter list, intuitively pick out a natural category of objects. But how are we to articulate what 'naturalness', in this context, amounts to?

One strategy for answering the question is to focus on what, if anything, is common to all the members of the class, and, relatedly, on the extent to which use of a given kind term delivers inductive or explanatory success. Nargs (which of course include buses, elephants and office blocks) have very little in common with one another; correspondingly, predictively speaking *narg* is an utterly useless kind, since nothing, aside from a lower bound on size, can be inferred from something's being a member of that kind. W. V. Quine (1969) claimed that kindhood and similarity are 'variations or adaptations of a single notion' (1969: 7), but that there are 'theoretical standards' of similarity that lend themselves to better inductions. Thus '[b]y primitive standards the marsupial mouse is more similar to the ordinary mouse than to the kangaroo; by theoretical standards the reverse is true' (1969: 15). Similarly, colour is 'king in our innate quality space, but undistinguished in cosmic circles. Cosmically, colours would not qualify as kinds' (1969: 14). So we might attempt to characterize the *natural* kinds as those kinds (unlike *red thing* and *mouse*, where the kangaroo mouse counts as a mouse) that align with 'theoretical standards' of similarity, thereby delivering superior inductive inferences. (This is not Quine's way, however. Quine holds that it is 'a very special mark of the maturity of a branch of science that it no longer needs an irreducible notion of similarity and kind' [1969: 52].)

Predictive (and explanatory) success has played a large role in the conception of natural kinds adopted by many philosophers of science (see e.g. Griffiths 1999 and Boyd, this volume), and in particular in Richard Boyd's 'homeostatic property cluster' account (1991). Many metaphysicians, however, hold that there must be something metaphysically or ontologically distinctive about natural as opposed to non-natural kinds; for example, that they are universals in something like David Armstrong's (1978) sense (see e.g. Ellis 2001: 67–8), or that they have essences in a metaphysically substantive sense (about which more later). For such metaphysicians, predictive success—while it might be a common or perhaps even a universal feature of natural kind concepts—cannot function as their defining characteristic, because such success comes in degrees. As Quine notes, '[b]etween an innate concept of similarity or spacing of qualities and a scientifically sophisticated one, there are all gradations. Science, after all, differs from common sense only in degree of methodological sophistication' (1969: 15). There is no cut-off, predictively speaking, anywhere in the spectrum between physics, chemistry, biology, the social sciences, and our ordinary, commonsense worldview; thus those who seek to endow natural kinds with a special metaphysical status must look elsewhere for the determining features of natural kinds.

One approach to this issue, taken by Brian Ellis (2001), is to lay down apparently *a priori* criteria that determine whether a given kind is natural. Ellis lists six conditions, individually necessary and jointly sufficient, which any natural kind must satisfy (2001: 19–21). For some kind *K* to qualify as a natural kind it must (1) be objective (i.e. mind-independent), (2) be categorically distinct, having no ontologically vague boundaries, (3) be demarcated from all other kinds via its intrinsic properties, (4) allow for species variation, where, for instance, two isotopes are members of the same element-kind in virtue of possessing the relevant essence (atomic number), while nevertheless differing from one another intrinsically by having a distinct atomic mass, (5) form a species-to-genus hierarchy in cases where a particular is a member of two (or more) natural kinds, and (6) have an (intrinsic) essence that is both necessary and sufficient for kind membership.

Any argument for such criteria must presumably be *a priori*: if we were to look to the sciences to tell us what the criteria for natural kindhood are, we would surely conclude that, for example, biological species are natural kinds, and we would therefore have no reason to expect categorical distinctness (criterion 2) to hold, and may have to abandon intrinsicness as well (criterion 6). By contrast, Ellis rules out biological species precisely on the grounds that they violate categorical distinctness. Moreover, as Emma Tobin argues in her chapter, it is unclear whether chemical kinds—Ellis's paradigm natural kinds—meet his own hierarchy requirement (criterion 5). *Protein* and *enzyme* would seem to be perfectly good natural kinds: they 'delineate real boundaries in biochemistry' (this volume, **XX**). But they violate the hierarchy requirement: neither is a subcategory of the other. Again, it seems that Ellis will have to simply flat out deny that such kinds are genuine natural kinds.

From a less metaphysical perspective, however, there need be nothing wrong with, for example, counting *protein* and *enzyme* as natural kinds. If natural kindhood is determined by the explanatory and predictive role the relevant kind concepts play in scientific theorizing, there is no reason why we should expect all of the kinds, reference to which underpins the explanatory and predictive success of chemistry, to form a hierarchical structure. Similarly, from this perspective there need be nothing wrong with natural kinds that have vague boundaries: that there is no fact of the matter about, say, which creature was the first member of any given biological species does not in the least undermine, or give us any reason to doubt, the explanatory and predictive power of evolutionary biology.

Another approach to answering the question of what distinguishes the natural from the non-natural kinds focuses on semantic differences between the two kinds of term. As a first pass, one might attempt to appeal to the notion of semantic simplicity: 'cat' is a semantically simple term, whereas 'object bigger than a car' is semantically complex. However, this answer will not work for obvious reasons: we could easily invent a word ('narg', say) to pick out all and only objects that are bigger than a car. 'Narg' is not semantically complex, but it picks out exactly the same class of objects as 'object bigger than a car'. So if the latter term intuitively fails to pick out a natural category, so does the former. Similarly, expressions such as 'H_2O' and 'the element with atomic number 79' intuitively pick out natural kinds (at least they do if 'water' and 'gold' do), but they are not semantically simple.

A more sophisticated, and apparently more promising, appeal to the semantic features of kind terms—and one that has played a large role in the literature—comes from the causal theory of reference as applied to kind terms, articulated by Saul Kripke (1980, first published in 1972) and Hilary

Putnam (1975). Kripke's basic idea is that there is a semantically distinctive species of general terms—the natural kind terms—that is analogous to a semantically distinctive category of singular terms, viz., proper names.

On the face of it, there is no particular reason to expect there to be a deep connection between the metaphysician's concern to delineate a metaphysically distinctive category of natural kinds—those kinds that carve nature at its joints—and the philosopher of language's concern to delineate a *semantically* distinctive category of natural kind *terms*. While some of Kripke's own examples of natural kind terms—'gold', 'water', 'tiger'—might suggest that the metaphysician's natural kinds just *are* those kinds that are picked out by the semanticist's natural kind terms, it is worth noting that Kripke has other examples of alleged natural kind terms that are not standardly thought of as picking out natural kinds in the metaphysician's sense ('heat', 'sound', and 'lightning'—see Kripke 1980: 134). Moreover, not all metaphysicians hold that even Kripke's central examples of natural kind terms all correspond to genuine natural kinds in the metaphysician's sense: Ellis, in particular, holds that biological species (*tiger*, for example) are not natural kinds since members of a single species differ from each other intrinsically, making species ontologically vague. (One might have concerns about water too, given that what we refer to as 'water' virtually always contains some impurities; see e.g. Abbott 1997 and LaPorte 1998 for discussion). This suggests that there is at best some overlap between the semantic category of 'natural kind terms' and the metaphysical category of natural kinds that carve nature at its joints, or have explanatory value, or whatever.

Be that as it may, many metaphysicians have taken it for granted that the natural kinds, in the metaphysician's sense, just *are* those kinds that are picked out by Kripkean natural kind terms. A large part of the motivation for this assumption is that the semantically distinctive category of natural kind terms promises to generate truths—what Kripke calls 'theoretical identities'—such as 'gold is the element with atomic number 79' that are metaphysically necessary but knowable only *a posteriori*.

We shall return to the (alleged) significance to metaphysics of the possibility of necessary *a posteriori* truths about natural kinds in §3 following, after introducing, in §2, Kripkean semantics and discussing some of the issues arising from it, as they have played out in the literature, including the first few chapters of this volume.

2. NATURAL KIND SEMANTICS

2.1 Kripke on proper names

According to the Frege-Russell view of names, names have both denotation and connotation. A name like 'Marie Curie' has both a referent, Marie Curie, and something 'besides that to which the sign refers . . . where in the mode of presentation is contained' (Frege 1892: 210). The mode of presentation, or *sense*, which determines the referent of the name, is usually taken to be some simple description. The sense of 'Marie Curie' might be 'the first woman to be awarded the Nobel Prize', for example. The motivation behind the sense/reference distinction is that it offers a solution to what has become known as 'Frege's Puzzle': the cognitive asymmetry between, for example, (a) 'Marie Curie is Marie Curie', which is both necessary and knowable *a priori*, and (b) 'Marie Curie is Maria Sklodowska', which appears to be *a posteriori*, and a useful extension of our knowledge. By distinguishing sense from reference, Frege dissolves the puzzle: although both 'Marie Curie' and 'Maria Sklodowska' have the same reference or denotation, they differ in sense (connotation), and it is connotation that determines the cognitive content of a sentence. More simply, the claim is that while the two names may refer to the same person, they differ in meaning, and this explains the cognitive asymmetry between (a) and (b).

Kripke famously objected to the Frege-Russell view of names, offering three different arguments, often referred to as the 'modal', 'epistemological', and 'semantical' arguments (Salmon 2005: 23–31). The modal argument shows that there are different truth conditions for sentences containing names and sentences containing descriptions, when assessed at different possible worlds. For instance, if 'Marie Curie' is synonymous with the description 'the first woman to win the Nobel Prize', then just as (m1) 'Marie Curie is Marie Curie' is necessary, so too is (m2) 'Marie Curie is the first woman to win the Nobel Prize'. But since there is a possible world, w_1, where there are no Nobel Prizes but where Marie Curie exists, (m1) will be true at w_1, while (m2) will be false. Hence (m1) cannot express the same proposition as (m2), and the description 'the first woman to win the Nobel Prize' cannot give the meaning of the name 'Marie Curie'.

The epistemological argument, like the modal argument, shows that there are different truth conditions for sentences containing names and those containing descriptions. However, the focus of the second argument is epistemic. Contrast (e1) 'if Marie Curie exists, then Marie Curie is Marie Curie', which is knowable *a priori*, with (e2) 'if the first woman to win the Nobel Prize exists, then the first woman to win the Nobel Prize is Marie Curie', which is clearly knowable only *a posteriori*. The Frege-Russell view entails, however, that the epistemological status of (e1) is the same as (e2)—both are knowable *a priori*. But consider a different possible world, w_2, where Irène Joliot-Curie is the first woman to win the Nobel Prize. Since this possibility cannot be ruled out *a priori*, (e2) cannot possibly be knowable *a priori*. Analysing the meaning of 'Marie Curie' as 'the first woman to win the Nobel Prize' leads us to misclassify (e2) as knowable *a priori* when it is in fact *a posteriori*. Hence the meaning of a name cannot be such a description.

Finally the semantical argument shows that descriptions always underdetermine the reference of names, even at the actual world. For instance, it is surely a coherent epistemic possibility that the history books got it wrong, and in fact Irène Joliot-Curie was the first woman to win the Nobel Prize. In that case, the description 'the first woman to win the Nobel Prize' in fact denotes Irène, rather than Marie. But since we have analysed the meaning of 'Marie Curie' as 'the first woman to win the Nobel Prize', it turns out that 'Marie Curie' in fact referred to Irène Joliot-Curie all along. Since we surely do not want to say that we intended to refer to Irène all along—perhaps we have never even heard of her until now—it seems clear the description 'the first woman to win the Nobel Prize' does not determine who the referent of 'Marie Curie' is. Hence reference to Marie Curie is underdetermined by description.

Kripke's positive proposal is broadly Millian: proper names have 'denotation' but lack 'connotation', to use Mill's terms, or (in Frege's terms) they have reference but no sense. Names 'do not refer to their referent by specifying a condition which their referent uniquely satisfies' (Hughes 2004: 1) but rather refer directly. The metasemantic part of the story is that there is a 'name-acquiring transaction' (Evans 2008: 316) where some object, *x*, is picked out by a language user in an 'initial baptism which is explained in terms either of fixing a reference by a description, or ostension' (Kripke 1980: 97). (In the case of reference-fixing by description, however, the description is not synonymous with the name. If I say 'let the caped crusader be called "Batman" ', my use of the expression 'the caped crusader' merely fixes the referent of 'Batman' as *that* person. It does not thereby preclude the possibility that the person so identified might stop wearing the cape or give up crusading in favour of organized crime.) Reference to the object (in this case, Batman) is then maintained by a causal chain of what Evans calls '*reference-preserving* links' (Evans 2008: 316), where speakers who are causally downwind of the baptism intend their use of the name to refer to whatever was initially referred to at the name-acquiring transaction. As a corollary of being non-descriptive, names get to be what Kripke calls *rigid designators*, which is to say that 'in every possible world it [the name] designates the same object' (Kripke 1980: 48).

Consider the name 'Ehrich Weiss'. Presumably when Ehrich was born his parents performed something like Kripke's initial baptism—a name-acquiring transaction—and dubbed him 'Ehrich Weiss'. Since proper names are rigid designators—they refer to the same object in all possible worlds—we can say that Ehrich's parents *stipulate* that the name 'Ehrich Weiss' is to apply to him, and this stipulation allows the name 'Ehrich Weiss' to denote Ehrich across all possible worlds. The upshot is we can ask counterfactual questions about *him* without worrying about being able to identify him in some possible world. As Kripke claims, using Nixon as his example:

> [A]lthough the man (Nixon) might not have been the President, it is not the case that he might not have been Nixon (though he might not have been called 'Nixon') . . . [and] it is because we can refer (rigidly)

> to Nixon, and stipulate that we are speaking of what might have happened to him (under certain circumstances), that 'transworld identifications' are unproblematic in such cases. (Kripke 1980: 49)

In some cases, however, individuals bear more than one name. Consider, again, Ehrich Weiss. At some point during his life, when he became a professional magician, Ehrich took on the new name 'Harry Houdini'. Given the Kripkean model, we can say that an additional name-acquiring transaction takes place, and Ehrich is now the referent of both the name 'Harry Houdini' and the name 'Ehrich Weiss'. Each independent baptism is the source of its own causal chain of reference-preserving links, making it entirely possible that different speakers can both be referring to the very same person using different names. Thus imagine two normal language users, Lu_1 and Lu_2. Lu_1 uses the name 'Ehrich Weiss' to denote Ehrich, and Lu_2 uses the name 'Harry Houdini' to denote Ehrich, yet neither speaker believes that they are denoting the same person as the other is denoting. Nevertheless, on Kripke's model, since both 'Ehrich Weiss' and 'Harry Houdini' are rigid designators, and the identity sentence 'Harry Houdini is Ehrich Weiss' is true, it follows that it is also necessary. More schematically, given that proper names are rigid designators, if '*a*' and '*b*' are co-referential proper names, and the identity sentence '*a* is *b*' is true at the actual world, then it is true in all possible worlds.

Kripke's conclusion draws an important distinction between two types of necessity: metaphysical and epistemic. The identity sentence 'Harry Houdini is Ehrich Weiss' is *metaphysically* necessary given (i) the necessity of identity, (ii) the rigidity of proper names, and (iii) the truth of the identity sentence. However, the identity claim is not *epistemically* necessary since its truth does not follow merely from reflection on the names 'Harry Houdini' and 'Ehrich Weiss'. There is a contrast between the identity sentence 'Ehrich Weiss is Harry Houdini' and 'a bachelor is an unmarried man'. Although both are metaphysically necessary, only the latter is epistemically necessary: its truth follows merely from reflection on the meanings of the terms 'bachelor' and 'unmarried man', and as such it is knowable *a priori*. The former, on the other hand, lacks epistemic necessity, and is thus knowable only *a posteriori*.

2.2 The Extension to Natural Kind Terms

With his account of the semantics of proper names in place, and the category of the necessary *a posteriori* established, Kripke goes on to extend his account to the 'more complex and philosophically significant case' of natural kind terms (Soames 2002: 242). According to Kripke:

> my argument implicitly concludes that certain general terms, those for natural kinds, have a greater kinship with proper names than is

> generally realized. This conclusion holds for certain for various species names, whether they are count nouns, such as 'cat', 'tiger', 'chunk of gold' or mass terms such as 'gold', 'water', 'iron pyrites'. (1980: 134)

Precisely what this kinship is intended to be is perhaps unclear, but at first blush we might expect general terms to acquire their reference via an initial baptism, which is then maintained by causal reference-preserving links, to be non-descriptive, and to be rigid designators. Less clear, however, is exactly why Kripke thinks that so-called 'theoretical identity sentences' containing natural kind terms, such as 'water is H_2O' and 'gold is the element with atomic number 79', are (like 'Harry Houdini is Ehrich Weiss') metaphysically necessary yet knowable *a posteriori*.

Theoretical identities are different to standard identities in two obvious respects. First, the phrases on the right-hand side of Kripke's paradigmatic theoretical identities—such as 'the element with atomic number 79' and 'H_2O'—are semantically complex, and do not obviously resemble proper names such as 'Ehrich Weiss'. Second, Kripke's discussion of theoretical identities is motivated, in part, by his observation that 'science attempts, by investigating basic structural traits, to find the nature, and thus the essence (in the philosophical sense) of the kind' (1980: 138), where these discoveries are expressed by theoretical identities; they are, as Kripke says, 'identity statements of essence' (1980: 159). Thus, as Salmon (2005: 82) notes, there are two distinct categories of necessary *a posteriori* truth in Kripke. The first, familiar category already discussed includes examples like 'Ehrich Weiss is Harry Houdini'. Their necessity and *a posteriority* is a product of (i) philosophical semantics, (ii) the necessity of identity, (iii) the substitution of two co-referring terms, and (iv) an uncontroversial empirical observation. The second, which includes theoretical identities and the necessity of origin thesis, appear to rely on a nontrivial notion of essentialism: their *a posteriority* is still the product of uncontroversial empirical observation, but their necessity stems from our beliefs about [our beliefs about? is that really right?] the properties of substances and objects. [moved from a bit later:] As Joseph LaPorte notes, theoretical identities 'resemble "Cicero is the product of the egg and sperm that generated him" rather than "Cicero = Tully" in respect that they aim to expose essence' (LaPorte 2004: 37).

Following convention let us label the mechanism for generating *a posteriori* necessities concerning natural kinds—which belong in the second category mentioned earlier—the *OK*-mechanism (Salmon 2005: 166). The argument employed to derive the necessity of, for instance, 'tungsten is the element with atomic number 74', looks something like the following:

1. It is necessarily the case that: something is a sample of tungsten if and only if it is a sample of *dthat* (the same substance as *this* is a sample of).
2. *This* (substance sample) has the atomic number 74.

3. Being a sample of the same simple substance consists in having the same atomic number.

Therefore,

3. It is necessarily the case that: every sample of tungsten has atomic number 74.

As Salmon's analysis illustrates, (1) is a consequence of Kripke's semantic thesis about the introduction of names and natural kind terms, (2) is an uncontroversial empirical observation, and (4) is the necessary *a posteriori* conclusion. Premise (3), however, is a nontrivial essentialist premise about the nature of substances that is independent of theories of meaning and reference. Salmon's own conclusion, as is well known, is that since premise (3), or some variant thereof, is required to derived *a posteriori* necessities for natural kinds as per Kripke's thesis in *Naming and Necessity*, *a posteriori* necessities involving natural kind terms are independently controversial, and hence the extension from names to natural kind terms is, in that respect, unsuccessful.

Salmon's analysis has led to a broad division in attempts to extend Kripke's insights from proper names to natural kind terms, much of which has focused on the extension of the notion of rigidity. On the one hand there is Scott Soames's claim that the extension of Kripke's insight is an important 'piece of unfinished [semantic] business' (Soames 2002: 242) left to us by *Naming and Necessity*. Soames's attempt to finish the business treats the relevant terms as predicates (' . . . is water', ' . . . is H_2O'), and theoretical identity statements as universal generalizations that turn out, *a posteriori*, to be necessarily true. On the other hand, Salmon proposes that we treat the terms on *both* sides of a theoretical identity statement as names of kinds, thus maintaining an 'imposing analogy' (1982: 43) between proper names and natural kind terms.

Soames's (2002) proposal, then, treats natural kind terms as predicates rather than common names, and analyses theoretical identities as universally quantified conditionals and bi-conditionals, so that, for example, 'necessarily, water is H_2O' is analysed as 'necessarily, x is water iff x is H_2O'. Soames recognizes a distinction between simple natural kind predicates like 'is water', and complex natural kind predicates like 'is H_2O'. Simple natural kind predicates are analogous to proper names insofar as they are non-descriptive, directly referential, and therefore rigid. Despite being predicates, natural kind terms do not denote their extensions (e.g. all *samples* of water), but rather denote natural kinds themselves (which Soames treats as intensions, that is, functions from worlds to extensions). Complex natural kind terms are also predicates, but in contrast to simple natural kind terms their reference is mediated rather than direct: these complex predicates are 'analogous to singular definite descriptions' (Soames 2002: 279), and their reference is fixed not by stipulation, but by scientific discovery. Complex natural kind predicates describe, and denote, kind-determining properties: in the water case, then, 'H_2O' denotes the kind-determining property of *being a substance, molecules of which contain two hydrogen atoms and one oxygen atom*. The *a posteriority* of '*x* is water iff *x* is H_2O' is a product of the different meanings of the respective terms. The necessity, however, stems from the fact that it is 'a feature of any genuine substance *S* that whatever its molecular structure turns out to be, all possible instances of *S* share that structure (and all possible instances of that structure are instances of *S*)' (Soames 2002: 273).

Soames claims that theoretical identities are 'linguistically guaranteed to be necessary if true' (2002: 279). However, this claim has come under some criticism, since, as we have seen, the guarantee of necessity seems to rely implicitly on a nontrivial essentialist assumption (Salmon 2003: 489). For instance, from the fact that (i) 'water' is a successfully introduced natural kind term and (ii) the sample(s) of water with which we are acquainted have the molecular structure H_2O, it does not follow that, *necessarily*, water is H_2O. It is only once we include claim that all possible instances of *S* share such-and-such a structure, and that all instances of that structure are instances of *S*, that the necessity of the 'identity' is guaranteed: this, in effect, is the *OK*-mechanism, sketched earlier. If this criticism is correct, Soames's extension of Kripke's semantics is inconsistent with the rejection of essentialism. Furthermore, it falls short of its goal of extending a purely linguistic mechanism for generating *a posteriori* necessities, since (if Salmon is right) it rests firmly on an assumption about the necessity of certain properties for particular kinds.

Salmon's alternative proposal is an explicit attempt to bypass the controversy over the essentialist premise. He claims that all general terms are analogous to logically proper names; thus his account is a genuine extension of Kripkean semantics for proper names to the natural kind case. According to Salmon, then, phrases like 'the element with atomic number 74' and chemical formulae like 'H_2O' are just the 'general-term version of a proper name whose reference is fixed through scientific convention concerning chemical-compound terms' (2003: 48). When 'H_2O' appears in a theoretical identity, then, it functions in much the same way as the proper name 'R2-D2' does in a standard identity sentence (1987/8: 197, n5; 2003: 488). The theoretical identity 'water is H_2O' is precisely analogous to 'Marie Curie is Maria Sklodowska' in that both contain two co-referring terms, derive their necessity via the simple substitution of terms, and are *a posteriori* since the two rigid general terms have independent chains of reference that lead back to two independent name-acquiring transactions, both of which denote a single referent. The difference, of course, is that where logically proper names denote individuals, general terms denote kinds, where kinds are construed as abstract entities of some sort.

The most significant criticism that has been raised to Salmon's proposed extension is that it trivializes the notion of rigidity, since *all* general terms—including descriptive general terms—turn out to be rigid (see Soames 2002 and Schwartz 2002). If we treat general terms as the names for kinds, then just as 'water' rigidly refers to the water-kind, so 'fridge' refers to the refrigerator-kind. Similarly descriptions such as 'the transparent, potable liquid that flows in the lakes and rivers' will pick out the same property in all possible worlds, namely the property of being transparent and potable and flowing in the lakes and rivers. Thus Kripke's insight that there is an asymmetry between natural kind terms and other general terms, in that only the former are rigid, is not maintained by Salmon's extension of Kripke's semantics for proper names.

In this volume Corine Besson offers a solution to the trivialisation problem, and the apparent symmetry between the semantics of natural kinds terms and other general terms. Besson's solution distinguishes between different notions of rigidity, and argues that only natural kind terms are *de jure* obstinately rigid designators. That is to say, natural kind terms designate their designatum across *all* possible worlds (obstinate rigidity), rather than some restricted set of worlds where the designatum exists, and their obstinate rigidity is a matter of stipulation (*de jure*). The asymmetry between natural kind terms and other general terms is grounded in a difference in metasemantics: natural kind terms, unlike other general terms, are introduced via name-acquiring transactions, or 'dubbings' as Besson (following Kaplan) calls them, and must satisfy some clearly defined metasemantic conditions in order to qualify as natural kind terms. These conditions, however, cannot be satisfied by other general terms, in particular descriptive ones, making natural kind terms unique among general terms.

Genoveva Martí and José Martínez-Fernández defend the view that general terms designate abstract entities (*The View*, to use their terminology) against the charge that it cannot provide an adequate account of rigidity for general terms. They identify and address three distinct objections: (a) the trivialisation problem, (b) the 'overgeneralization problem' (natural kind terms are rigid, but other non-descriptive general terms, such as 'bachelor' and 'pencil', are non-rigid), and (iii) the claim that 'The View' cannot provide an account of the necessity of true theoretical identities.

First, they address the trivialisation problem, a problem that can be stated in terms of truth conditions: when a general term is used as a predicate there seems to be no difference in contribution to truth conditions between rigid and non-rigid readings of a given general term. They respond by finding a case where there is, in fact, a difference in truth conditions depending on whether the relevant predicate is treated as rigid or as non-rigid.

Second, they tackle the problem of overgeneralization: what we earlier referred to as the symmetry of general terms. If 'water' designates the water-kind, then it looks like 'fridge' designates the refrigerator-kind, and where the former is rigid so too is the latter. Hence there is no asymmetry between natural and non-natural kind terms. Here, Martí and Martínez-Fernández bite the bullet, claiming that the symmetry of general terms is as it should be, since the semantic behaviour of all simple general terms is the same: there is just no interesting semantic distinction to be made between, say, 'philosopher' on the one hand, and 'tungsten' on the other.

Finally they address the problem of deriving the necessity of theoretical identities. Here they reject Soames's view that theoretical identities are universally quantified bi-conditionals, observing that theoretical identities are claims about kinds, not about the objects that fall under the extension of kind-predicates. By distinguishing between the predicative and denoting use of a general term—*exemplifier* and *kind* semantics, to use Martí and Martínez-Fernández's terminology—they offer an account of why true theoretical identities are necessary.

Åsa Wikforss, in contrast to Besson, argues against the treatment of natural kind terms as a special semantic category of general term. According to Wikforss, not only should the semantic difficulties associated with natural kind terms be taken seriously, but discussion of natural kind terms is intimately connected to talk of natural kinds and natural kind essentialism, since the former are taken to designate the latter. But it is highly implausible to suppose that 'there is a category of terms that is special from a semantic point of view, even though identifying this category depends on the development of sophisticated empirical theories, such as contemporary chemistry or evolutionary theory' (this volume, **XX**). Wikforss concludes that the Kripke-Putnam account of natural kind terms should be rejected in favour of a version of descriptivism, where the meanings of natural kind terms are given by sets of properties, including properties describing the law-like behaviour of the kind. There is, then, no distinctive *semantic* feature of natural kind terms compared to other general terms. Any special significance they have is a product of the role they play in scientific explanation.

Harold Noonan is also unconvinced by the Kripkean view of natural kind terms, and his chapter seeks to explain the rigidity of both proper names and natural kind terms in a way that is consistent with the Frege-Russell account. However, Noonan dismisses the treatment of theoretical identities as revealing truths that are metaphysically necessary yet knowable *a posteriori*, arguing that the Kripke's category of the necessary *a posteriori* in the case of natural kinds stems from the conflation of the notion of an essential property, and an *a posteriori* but only contingently true thought. For Noonan, Kripkean metaphysical necessity belongs on Hume's bonfire.

As the foregoing discussion illustrates, the Kripke-Putnam approach to natural kind terms remains highly controversial. Does Kripke really refute the Frege-Russell view, or (as Noonan argues) can the semantic facts to be explained be accommodated by a broadly descriptivist account of kind terms? Is there even a nontrivial semantic phenomenon of rigidity in the case of kind terms? And, if so, to what extent does the phenomenon apply to terms that denote genuine natural kinds in the metaphysician's and philosopher of science's sense?

3. KRIPKE, *A POSTERIORI* NECESSITY, AND THE METAPHYSICS OF NATURAL KINDS

As we have seen, some philosophers of language hold that the prospects for a category of general terms that both is semantically distinctive and has as its members general terms that pick out genuine natural kinds in the metaphysician's and philosopher of science's sense are decidedly bleak. Be that as it may, many metaphysicians have tied their metaphysical views to Kripke-Putnam semantics—in other words, they have taken it for granted that the class of Kripkean 'natural kind terms' will pick out all and only the natural kinds in the metaphysician's sense.

What is the explanation for this desire to align Kripke-Putnam semantics with (what one might think is) an independent account of the metaphysics of natural kinds? Well, in fact there are two distinct (though perhaps related) strands in metaphysics that need to be disentangled: one allied to Kripke's concerns, and in particular to the alleged necessary *a posteriori* status of 'theoretical identifications', and one more naturally allied to Putnam, whose prime concern was to ward off the threat of Kuhnian incommensurability. In this section we discuss the Kripkean lineage, returning to Putnam in §4 following.

The strand of metaphysics that is most closely allied to Kripke's concerns takes its cue from Kripke's claim that theoretical identities, such as 'water is H_2O' and 'gold is the element with atomic number 79', are metaphysically necessary but knowable only *a posteriori*. This is a claim that apparently promises to allow metaphysicians to resurrect the largely discarded Aristotelian notion of *essence*, but this time shorn of the commitment to rationalism that had been spurned by the empiricist tradition.

Consider, by way of a foil, Descartes' well-known discussion of wax in his second *Meditation*. According to Descartes, the intellect—not the senses, and not the imagination—grasps what the wax *is*: 'what is left' when 'we take away everything which does not belong to the wax' (1996: 20). In other words, the essence of wax (it is extended, flexible, and mutable) is grasped by the intellect, with no help from the senses. Knowledge of wax's essence is thus *a priori*, in the sense that no possible sensory experience could show my (clear and distinct) idea of it to be mistaken. I may be mistaken about wax's existence, but not its essential nature.

There are several relevant strands to the traditional empiricist line on essences. Locke was scathing about Aristotelian essences. Firstly, as an empiricist, he objected to Aristotle's claim that we can come to have *a priori* knowledge of the essences of substances: the investigation of substances is a purely *a posteriori* matter. Secondly, he denied the ontology of universals which was integral to Aristotle's notion of essence, claiming that 'General and Universal, belong not to the real Existences of Things' (1692: III.iii.11), and that 'all things that exist are only particulars' (1692: III.iii.6). For Locke, Aristotelians made the mistake of thinking that their theory of essences hooked onto bona fide distinctions in nature by treating 'distinctions in thought for real distinctions, abstractions for realities' (1692: III. vi. 10). Locke's own, alternative view was that we observe the manifest similarities between discrete particulars, and via abstraction formulate the notion of a (natural) kind. We then use this abstract idea, or 'nominal essence', to classify discrete particulars that correspond to our idea as members of the kind. Once these manifest properties have been singled out as defining characteristic of the kind, we can, in principle, cast around for the underlying properties that explain why *particular* members of the kind have the manifest properties they do. These underlying features might vary between members of a given kind; just as watches vary enormously in the underlying mechanisms by virtue of which they get to tell the time (and hence count as watches), so too might individual samples of gold or water.

In the case of substances, the prospects of finding out which underlying properties explain the manifest ones were, in Locke's own time, bleak. The important point for Locke, however, was that the 'real essence' would explain the 'nominal essence', but since the 'nominal essence' was our abstraction, the 'real essence' too would be conventional in this sense. Again, the real essence of a particular watch (and perhaps of samples of gold and water too) is an abstraction from its full underlying nature, since some of its underlying features play no role in its having the characteristics definitive of watchhood. For Locke, then, what for Descartes is the intellectual grasp of the essence of wax is no more than a semantic decision about the characteristics of wax that constitute its nominal essence. Wax doubtless has underlying features that could in principle explain its malleability and so on, but these features are unknowable (in practice) and therefore irrelevant to its correct classification as wax. (See Leary 2009 for a fuller discussion.)

Hume agreed with Locke that we cannot 'penetrate into the essences' of things; but, his concern being causal reasoning rather than classification, his aim was to show that we cannot do this in such a way as to gain rational insight into the causal structure of the world: '[i]f we reason *a priori*, anything may appear able to produce anything' (Hume 1748: 164). Thus, since knowledge of essences—which would deliver knowledge of how things *must* behave—would have to be *a priori* knowledge, metaphysical speculation about the existence and nature of essences belongs on Hume's bonfire: consigned to the flames on the grounds that it contains nothing but 'sophistry and illusion' (Hume 1748: 165). Thus on Hume's view, Descartes is simply wrong in claiming that wax's dispositional properties (flexibility and mutability) are knowable *a priori*; it is perfectly conceivable, if we 'consider the matter *a priori*', that a piece of wax should retain its shape no matter what we do to it.

One (doubtless controversial) way to cast the issue runs as follows. For both Locke and Hume—as empiricists—what was fundamentally wrong with essentialism was the idea that there is some special rational faculty that could grasp the essences of things. In other words, it was the alleged epistemology of essences that was primarily to blame for the dubious metaphysics: the oddity of essences—their being features of objects that are somehow graspable by the intellect—was a consequence of the need for *a priori* epistemic access to the nature of reality. Kripke's argument for the existence of 'theoretical identity' claims that are metaphysically necessary but knowable only *a posteriori* thus delivered the prospect of a version of essentialism that renders essences unmysterious because they cannot be known *a priori*. Once we grant that it is empirical investigation that delivers knowledge of essences, essences no longer need to be seen as peculiar features that are somehow capable of being grasped by the intellect. (This, of course, is more rational reconstruction than historical fact.) Thus Brian Ellis, for example, presents his 'scientific essentialist' position as one that ought to be embraced by scientific realists, and that explains the explanatory and predictive success of the relevant sciences (see e.g. Ellis 2001: 1–2).

Moreover, the possibility that natural kinds, or the properties that characterize them, might turn out to have *dispositional* essences delivers the prospect of a similar manoeuvre in the causal case. For now, in principle, we can agree with Hume that we cannot tell *a priori* how a given object or substance will behave (that electrons are disposed to repel each other, say) without the need for spurning an 'anti-Humean' metaphysics, according to which behaviour and identity come apart. Rather than seeing the laws of nature as contingent generalizations about how objects and substances in fact behave, we can see them as characterizing the essences of properties or kinds, such that those properties or kinds would not be the properties or kinds that they are if they behaved differently. Thus we can in principle hold that the laws of nature are metaphysically necessary while denying that we can know them *a priori* (see Ellis 2001: 219–20 and Bird 2007: Ch. 8).

If the necessary *a posteriori* status of theoretical identities really can do the job of rendering essentialism 'scientific', then of course the scope of the necessary *a posteriori* needs to be broad enough to cover theoretical identities across all the sciences that deal in natural kinds: minimally physics and chemistry, and arguably also biology. This is an issue that is taken up in several of the chapters in this volume. Joseph LaPorte, building on his 2004 book, argues that natural kind essences are 'stipulated' rather than discovered, and so theoretical identities such as 'water is H_2O', while necessary, are not *a posteriori*. Investigation into the chemical constitution of water did not *reveal* that water is essentially H_2O; instead, it provoked a *decision* about what 'water' would refer to; in particular, the decision that D_2O or 'heavy water' (whose molecules have deuterium, an isotope of hydrogen, as a constituent) counts as water. Use of the term 'water' prior to this stipulation was 'open-textured' or vague: it was indeterminate whether samples of pure D_2O counted as water or not. A different decision could have been made without error: scientists visiting Deuterium Earth, where what flows in the rivers and lakes is D_2O (which is poisonous for many organisms, including humans), might well have decided that D_2O is not water, and hence that water is not essentially H_2O (since a sample can be H_2O without being water, by being a sample of D_2O). LaPorte argues that a similar story can be told for other theoretical identities, and in particular biological essence claims such as '*Mammalia* is the clade with stem *M*', where *M* is the nearest common ancestor of horses and echidnas.

Alexander Bird raises two objections to LaPorte's view. First, he argues that an alternative account can be given of the relevant phenomenon—what LaPorte sees as precisification—that upholds the discovery thesis. In the case of water, for example, LaPorte agrees with Bird that ' "from the point of view of chemistry" it would have been a mistake to say that D_2O is not water' (this volume, **XX**), but claims that 'with respect to a term like "water", *chemical* properties are not decisive and physical properties, including biological properties and physical properties, matter too' (**XX**). Thus it would not be a 'mistake' to rule out D_2O from counting as water on the basis of its failure to perform the life-supporting function that naturally occurring samples of water here on Earth actually perform. Bird disagrees. Starting from Linus Pauling's claim that '[c]hemistry is the science of substances', he argues that, given that water is a substance, it is chemical facts that determine *the* correct classification of water. Hence 'water' always has determinately included D_2O in its extension. Thus the *a posteriori* status of 'water is H_2O' is preserved.

Second, Bird argues that LaPorte's arguments have limited application. In particular, many natural kind terms will be neither vernacular (such as 'water' and 'rodent')—the case that LaPorte focuses on—nor stipulatively introduced (such as 'mendelevium is the element with atomic number 101'—more on this kind of case following). For example, 'actinium' was introduced as the name for an element at the turn of the twentieth century. However, the discovery of actinium—and the introduction of the name—predated Henry Moseley's introduction of the notion of atomic number, which 'tells us the nature of the element' (this volume, **XX**). So clearly 'actinium' was not (unlike 'mendelevium') stipulatively introduced, precisely to refer to the element with atomic number 89, since information about atomic number was not available at the time the term was introduced. However, 'actinium' clearly referred to the very same element at the turn of the twentieth century as it does now, and so there is simply no room in this story for anything like the precisification required by LaPorte's 'stipulation, not discovery' thesis: there was no indeterminacy, akin to the indeterminacy LaPorte claims to have been present prior to the precisification of 'water', to be resolved. Hence even if LaPorte is right about vernacular natural kind terms, the *a posteriori* status of 'actinium is the element with atomic number 89' and its ilk is preserved.

The historical context touched on by Bird is fleshed out in much more detail in Robin Hendry's chapter. Hendry focuses on the scientific context of Lavoisier's introduction of the terms 'oxygen' and 'hydrogen' in order to shed light on the dispute about the semantics of chemical-kind terms. In effect, Hendry argues that what goes for 'actinium' goes for 'oxygen' and 'hydrogen'—and also 'water'—too: 'oxygen' determinately referred to the element with atomic number 8, even though Lavoisier's discovery significantly predated the development of the periodic table and the introduction of atomic number as the defining characteristic of elements. Since—thanks to his adoption of what Hendry calls the 'core conception' of elements—Lavoisier used the names regardless of the state of chemical combination (e.g. oxygen is present in metallic oxides), those names 'must have tracked a nuclear property that is invariant across chemical change' (this volume, **XX**). And, of the two relevant candidates—atomic number or nuclear charge (number of protons) and atomic weight (determined by the number of protons and neutrons, averaged across a sample which may include different isotopes with different atomic weights)—the names must have tracked atomic number (nuclear charge), since this is what determines chemical behaviour. Hendry's argument that the 'core conception' ensures the determinacy of reference of Lavoisier's 'hydrogen' also plays a role in his rejection of LaPorte's claim that the reference of 'water' was indeterminate prior to the decision to include D_2O within its extension: '[i]t seems highly implausible that, before the twentieth century, chemists' usage of 'water' was indeterminate as to isotopic extension, while the names of water's elemental components were regimented' (this volume, **XX**).

Our own contribution to this volume focuses on the third kind of case identified by Bird—theoretical 'identities' such as 'mendelevium is the element with atomic number 101' and 'ununbium is the element with atomic number 112'—and on Ellis's and Bird's claim that, as with natural kind essences, statements of the 'dispositional essences' of properties are metaphysically necessary but knowable only *a posteriori*. We argue that there are clear counterexamples to the claim that *all* allegedly essence-exposing statements about natural kinds (and fundamental dispositional properties) are knowable only *a posteriori*. For example, the name 'ununbium' is derived from rules laid down by the International Union of Pure and Applied Chemistry concerning the names of newly discovered or created elements. It is thus analytic that ununbium is the element with atomic number 112. For somewhat different reasons, we argue that the claim that the essences of dispositional properties are *a posteriori* discoveries is also difficult to maintain.

As the debates in these chapters show, it is a highly controversial question whether, or to what extent, the Kripkean programme can be extended to cover all the natural kinds to be found in the sciences (our own view is that it cannot). Why does this issue matter? Well, the answer depends on what function the Kripkean programme is being asked to perform. LaPorte, and to some extent Hendry, is more interested in the issue of referential stability across theory-change than in the metaphysics of essence and necessity (see §4 following). From that perspective, it need not be fatal to a given view that some theoretical identities are knowable *a posteriori* and some are not. For example, LaPorte's view is that, given indeterminacy of prior usage, a causal theory of reference does not block referential instability. But that view could in principle hold only for a limited class of cases; it might be that rigid designation gets us *some* referential stability (e.g. for 'oxygen'), but just not as much as some philosophers have thought; for example, we don't get stability for 'water' (see LaPorte, this volume, n.8 for the explicit acceptance of the claim that what goes for 'water' does not generalize to all natural kind terms).

From the perspective of the kind of view articulated earlier in this section, however—according to which the *a posteriori* status of theoretical identities cuts essentialism loose from its association with the power of the intellect to penetrate the nature of the material world and renders it 'scientific'—the situation is rather more black-and-white. As we point out in §3 of our chapter, Ellis himself takes *a posteriori* necessity to be the defining feature of what he calls 'real' necessity, as opposed to the mere *de dicto* necessity that analyticity delivers. On this view, there is no room for a 'mixed' position; one can hardly take the necessity of 'gold is the element with atomic number 79' to be 'real' and the necessity of 'ununbium is the element with atomic number 112' to be any less real. The difference in the genesis of the respective terms, 'gold' and ununbium', reveals no underlying metaphysical difference: as far as the world is concerned, gold and ununbium are on a par.

This being so, one might wonder whether essentialism would be better off unshackling itself from controversial claims about the importance and extent of the necessary *a posteriori*. After all, the core of the essentialist position is a commitment to a particular view about modality: the view that modal facts flow from the intrinsic nature of the actual world (they are 'intra-world' facts) and not from facts about other possible worlds ('inter-world' facts). The latter view, popularized by David Lewis (1986a), is ideally suited to accommodate a Humean worldview, according to which all there is, at bottom, is a distribution of logically distinct 'local matters of particular fact, just one little thing and then another' 1986b: ix), which, in and of themselves, have no modal import. The former view, which Ellis himself upholds (see e.g. his 2001: 272–4), is shared by a number of essentialists who reject Ellis's claim that only *a posteriori* necessity is 'real' necessity (e.g. Oderberg 2007; Lowe 2008). Kit Fine (1994), for example, argues that analytic truths can be seen as articulating 'real definitions' of objects (and, presumably, kinds) and therefore as essence-revealing. And E. J. Lowe argues that 'knowledge of essence cannot be obtained from a combination of purely non-modal empirical knowledge—observational or experimental knowledge—and *a priori* logical or conceptual-cum-semantic knowledge' (2008: 25).

Whether or not essentialism thus conceived belongs on Hume's bonfire is a question we shall not explore here. One reason for Humeans to exercise caution with the firelighters, however, is articulated in Jessica Wilson's chapter. Wilson observes that in general Humeans—deniers of necessary connections between distinct existences, and specifically deniers of causal necessity—happily accept 'constitutional necessities', including, for example, the claim that water is necessarily constituted by H_2O molecules, or the claim that, necessarily, electrons are negatively charged. She argues, however, that 'if one accepts constitutional necessities, one should accept certain causal necessities' (this volume, **XX**). Non-Humeans can account for the metaphysical grounds of these necessities, and also explain how we have epistemic access to those grounds, by appealing to the 'modally stable' causal profiles of the relevant properties. Humeans, however, cannot appeal to modally stable causal profiles, since it is part and parcel of the Humean view that properties (or at least fundamental ones) do not have such profiles. To accept modal stability would be to accept that the causal role of a (fundamental) property fixes its identity, and this falls foul of what Wilson calls 'Hume's Dictum': no necessary connections between wholly distinct entities. In the absence of any alternative explanation of the relevant facts, Humeans violate the principle that we should 'accept the holding of those metaphysical facts that enter into the best account of the justificatory facts concerning claims we accept; hence if we accept certain constitutional necessities, we should accept certain causal necessities' (this volume, **XX**).

Before leaving the topic of essentialism, it is worth noting that 'essentialism' means different things to different authors. Previously, we characterized essentialism as a view about the sources of modality. However, many authors who discuss 'essentialism' clearly have quite different—and much weaker—theses in mind. Take, for example, the debate between LaPorte and John Dupré concerning whether the former's view constitutes a version of 'essentialism' (LaPorte thinks it is—see LaPorte, this volume, **XX**. Dupré thinks it isn't, or at least considers LaPorte's version of essentialism to be 'toothless'; see his 2004.) Dupré says:

> the possibility that scientists may construct necessary truths by defining their terms . . . says little of interest about essentialism. [LaPorte] is consistently opposed to the notion that essences are in any interesting sense discovered. Moreover, even though he thinks taxa have essences, for good biological reasons he does not believe that members of taxa

> necessarily share these essences. Biological essences that have nothing to say about individual organisms strike me as toothless. (Dupré 2004)

For Dupré, 'essentialism' is characterized by a commitment to the view that members of a natural kind share a 'real essence', where '[t]o assert that there are real essences is, in part, to claim that there are fundamental properties that determine the existence and extensions of kinds that instantiate them' (1986: 62). According to this characterization, LaPorte is no essentialist for two reasons. First, a natural kind may be defined by an extrinsic essence (as with cladism), which, in Dupré's terms, does not count as an essence at all (hence his remark that on LaPorte's view, taxa 'have essences' even though members of taxa need not 'share these essences': members of the clade *chordata* share the same stem, but of course this is a historical fact about them and not a matter of shared underlying intrinsic nature). Second, LaPorte shares Dupré's pluralist approach to natural kinds: in particular, both reject the idea that there is one, uniquely right, system of biological classification. For Dupré, this is sufficient to render LaPorte's view anti-essentialist, since '[a] pluralist, by denying that there is *any* uniquely correct scheme of classification, is clearly committed to the denial of essentialism' (1986: 53).

LaPorte, however, implicitly denies both of these alleged implications of 'essentialism'. (See Griffiths 1999 for an even more relaxed conception of 'essentialism' and a useful discussion of the role that concerns about essentialism have played in debates about biological classification.)

The nub of the dispute between LaPorte and Dupré, it seems to us, is really a matter of whether (to stick with biology for the moment), as Dupré maintains, the criteria that 'distinguish the members of a species [from members of other species] are likely to be chosen in part for anthropocentric reasons such as ease of human application' (1986: 36). More generally, Dupré's position is, as he describes it, 'promiscuous'; for example, '[i]f cedars . . . do all serve a common function for the carpenter, then relative to that human practice I see no reason to deny the naturalness of the kind they form: it is a natural fact, after all, that there exists a class of things, albeit botanically diverse, suited to this role' (1986: 63). LaPorte, by contrast, would not count cedar as a natural kind: the interests of carpenters are not the kinds of thing that can serve to make a kind 'natural'. If this is right, then the dispute between Dupré and LaPorte is not so much about essentialism as about the role of human interests in determining the naturalness of a kind.

4. PUTNAM, INCOMMENSURABILITY, AND NATURAL KINDS

At least one part of Putnam's aim in introducing his version of the causal theory of reference and the concomitant claim about the existence of necessary *a posteriori* truths about natural kinds was to provide an antidote to the threat of relativism posed by Kuhn's incommensurability thesis. On

Kuhn's view, 'after Copernicus, astronomers lived in a different world' (Kuhn 1996: 117); ' . . . after discovering oxygen Lavoisier worked in a different world' (1996: 118), and 'until that scholastic paradigm was invented, there were no pendulums, but only swinging stones, for the scientists to see. Pendulums were brought into existence by something very like a paradigm-induced gestalt switch' (1996: 120). In other words, the conceptual shifts involved in changes of paradigm are not merely reconceptualisations of the very same independently existing entities referred to by occupants of previous paradigms; the world that the occupants of the new paradigm inhabit and talk about is literally a different world.

Putnam (1975: 235–8) argues that his (and, by extension, Kripke's) account of *a posteriori* necessity undermines Kuhn's relativist position. It is not the case that for ordinary people prior to the relevant developments in chemistry, gold was a yellow, malleable metal, and then it 'became' the element with atomic number 79—or rather, some new entity, also called 'gold', came into existence. Rather, the substance referred to by the earlier speakers was exactly the same substance—gold—that later, chemically informed speakers referred to. As Putnam says:

> It is beyond question that scientists use terms as if the associated criteria were not *necessary and sufficient conditions*, but rather *approximately* correct characterizations of some world of theory-independent entities, and that they talk as if later theories in a mature science were, in general, *better* descriptions of the *same* entities that earlier theories referred to. In my opinion the hypothesis that this is *right* is the only hypothesis that can account for the communicability of scientific results, the closure of acceptable scientific theories under first-order logic, and many other features of the scientific method. (Putnam 1975: 237)

This concern with stability of reference across theory-change lies in the background of LaPorte's chapter. In his 2004, he argues that 'the causal theory does not live up to its promise':

> The causal theory of reference leaves room for plenty of reference change . . . causal baptisms, which according to the causal theory endow terms with their reference conditions, are performed by speakers whose conceptual development is not yet sophisticated enough to allow the speakers to coin a term in such a way as to preclude the possibility of open texture, or vague application not yet recognized. When speakers recognize that their use of a term is vague, they tend to offer further specification for its use. That further specification amounts to a stipulation that changes the term's meaning. (LaPorte 2004: 118)

As we have seen, Hendry argues that LaPorte's claim about stipulation does not hold in the case of 'oxygen', 'hydrogen', and 'water'; but he also takes issue with LaPorte's claim that Putnam's semantics is 'useless in blocking instability' (LaPorte 2004: 118). According to Hendry, Putnam's semantics is best seen as establishing only the *possibility* of referential stability across theory-change (a thesis that is still opposed to Kuhn's, since Kuhn appears to think that such stability is impossible). 'Referential stability (and, for that matter, referential instability or indeterminacy) can be inferred only once the actual historical details of the introduction of a particular kind-term are examined, along with the intellectual context in which it occurred' (this volume, **XX**). Hendry's basic criticism of LaPorte, then, is that he over-generalizes: LaPorte may well be right about, for example, elements discovered before about 1700 (e.g. gold and sulphur), since the discovery of these elements predates the 'core conception' of elements. But nothing follows, in the absence of the kinds of historical details Hendry provides in the case of 'oxygen', about other natural kind terms. It should be noted, however, that if LaPorte is guilty of over-generalization, plenty of philosophers of the opposite persuasion have over-generalized in the other direction, taking Putnam's work, just by itself, to have shown that referential instability is no longer a live issue (see LaPorte 2004: 117–8 and 117, n.8).

Putnam's concern with scientific realism is also taken up in Richard Boyd's chapter, which brings together much of his earlier work on naturalism, scientific realism, and natural kinds. For Boyd, the 'fundamental question which the theory of natural kinds addresses is this: "How do classificatory practices and their linguistic manifestations help to underwrite the reliability of scientific (and everyday) inductive/explanatory practices?" ' (this volume, **XX**). His answer involves the notion of 'accommodation': a fit between the inferential practices (involving use of the relevant kind terms) of a given discipline on the one hand, and causal structures in the world on the other. 'Definitions' of natural kinds are an *a posteriori* matter: subject to revision in the light of new empirical discoveries.

How 'realist' is Boyd's realism? On the one hand, for Boyd natural kinds are explicitly discipline-relative, so that, for example, someone who denies the reality of race can be understood as 'denying that races as currently understood, play an epistemically legitimate role in biology' (this volume, **XX**); but this does not preclude race being a natural kind in some branches of social science, for example, in the study of social stratification and inequality. He notes that 'the realist naturalist's conception of "reality" questions is less elevated than other conceptions might be, but that's the fate of naturalistic metaphysics' (**XX**). It is also explicitly anti-reductionist: we will find natural kinds wherever we have a discipline that satisfies what Boyd calls the 'epistemic access' and 'accommodation' conditions, whether the discipline is physics or pharmacology or a social science.

On the other hand, Boyd's conception of natural kinds is realist in the sense that natural kinds really do exist, and their existence is underpinned by real, mind-independent causal structures. His realism—in the sense of opposition to Kuhnian relativism—comes out in his discussion of conceptual role semantics (§5.4): *pace* conceptual role semantics, according to the accommodationist view alchemists succeeded in referring to sulphur and mercury despite the fact that the 'most central and explanatory patterns they associated with the relevant terms' were utterly misguided (**XX**). Similarly, human sociobiologists succeed in referring to real aspects of human behaviour—altruism, competition, and so on—despite the fact that sociobiology enshrines inferential commitments that are deeply confused.

Boyd's commitment to the Kripke-Putnam programme, then—or at any rate to a broadly causal theory of reference—is part of a broader commitment both to naturalism and to realism understood in (by many metaphysicians' standards) a somewhat deflationary sense. He is a realist about natural kinds in exactly, and only, the sense in which a philosopher of science might be a realist about molecules or species or natural selection. This approach stands in stark contrast to writers such as Bird (2007) and Ellis (2001), whose interest is in rescuing a robust form of essentialism from its rationalist roots.

As should by now be clear, the issues surrounding the semantics and metaphysics of natural kinds are many and complex, ranging from technical issues in the philosophy of language concerning the correct characterization of rigidity to very general questions concerning the tenability of scientific realism. We hope, in the course of this introduction, to have indirectly made the case that there is much scope for fruitful interaction between philosophy of language, metaphysics, and the philosophy of science on the topic of natural kinds, and to have highlighted some of the specific issues where such interaction might usefully take place.

REFERENCES

Abbott, B. (1997) 'A note on the nature of "water" ', *Mind*, 106: 311–9.

Armstrong, D. M. (1978) *Universals and Scientific Realism, Vol. II: A Theory of Universals*. Cambridge: Cambridge University Press.

Bird, A. (2007) *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.

Boyd, R. (1991) 'Realism, anti-foundationalism and the enthusiasm for natural kinds', *Philosophical Studies,* 61: 127–48.

Descartes, R. (2003) 'Meditations on first philosophy', in *Meditations on First Philosophy*, ed. and trans. J. Cottingham (2nd edn.), Cambridge: Cambridge University Press.

Ellis, B. (2001) *Scientific Essentialism*. Cambridge: Cambridge University Press.

Evans, G. (2008) 'The causal theory of names', in A. P. Martinich (ed.), *The Philosophy of Language*, New York: Oxford University Press, 314–25.

Fine, K. (1994) 'Essence and modality', *Philosophical Perspectives*, 8: 1–16.

Frege, G. (1892) 'Sense and reference', translated and reprinted in *The Philosophical Review*, 57: 209–30.

Griffiths, P. (1999) 'Squaring the circle: natural kinds with historical essences', in R. Wilson (ed.), *Species: New Interdisciplinary Essays*, Cambridge, MA: MIT Press, 219–28.

Hughes, C. (2004) *Kripke: Names, Necessity and Identity*. Oxford: Clarendon Press.

Hume, D. (1748) *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. L. A. Selby-Bigge, 3rd edn., revised and ed. P. H. Nidditch. Oxford: Clarendon Press 1975.

Kripke, S. (1980) *Naming and Necessity*. Oxford: Basil Blackwell. (First published in G. Harman and D. Davidson, eds., *Semantics of Natural Language,* Dordrecht, Netherlands: D. Reidel, 1972.)

Kuhn, T. S. (1996) *The Structure of Scientific Revolutions*, 3rd edn. Chicago: Chicago University Press. (First published 1962.)

LaPorte, J. (1994) *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press.

———. (1998) 'Living water', *Mind*, 107: 451–5.

Leary, N. (2007) 'How essentialists misunderstand Locke', *History of Philosophy Quarterly,* 26: 273–92.

Lewis, D. K. (1986a) *On the Plurality of Worlds*. Oxford: Blackwell.

———. (1986b) *Philosophical Papers, Vol. II*. New York: Oxford University Press.

Locke, J. (1689) *An Essay Concerning Human Understanding*. Fraser, ed. P. H. Nidditch, Oxford: Clarendon Press, 1979.

Lowe, E. J. (2008) 'Two notions of being: entity and essence', *Philosophy,* 23: 23–48.

Oderberg, D. S. (2007) *Real Essentialism*. Abingdon and New York: Routledge.

Putnam, H. (1975) 'The meaning of "meaning" ', in K. Gunderson (ed.), *Language, Mind and Knowledge*, *Minnesota Studies in the Philosophy of Science, VII*, Minneapolis: University of Minnesota Press. Reprinted in H. Putnam, *Mind, Language and Reality: Philosophical Papers, Volume II*, Cambridge: Cambridge University Press, 215–71.

Quine, W. V. (1969) 'Natural kinds', in N. Rescher (ed.), *Essays in Honor of Carl G. Hempel*, Dordrecht, Netherlands: D. Reidel, 5–23.

Salmon, N. (1987/8) 'How to measure the standard meter', *Proceedings of the Aristotelian* Society, 88: 193–217.

———. (2003) 'Naming, necessity and beyond', *Mind* 112: 475–92.

———. (2005) *Reference and Essence*, 2nd edn. New York: Prometheus Books.

Schwartz, S. (2002) 'Kinds, general terms, and rigidity', *Philosophical Studies*, 109: 265–77.

Soames, S. (2002) *Beyond Rigidity*: *The Unfinished Semantic Agenda of Naming and Necessity*. New York: Oxford University Press.

2 Rigidity, Natural Kind Terms, and Metasemantics

Corine Besson

1. INTRODUCTION: INITIAL DIFFICULTIES WITH RIGIDITY AND NATURAL KIND TERMS

A paradigmatic case of rigidity for singular terms is that of proper names. And it would seem that a paradigmatic case of rigidity for general terms is that of natural kind terms. However, many philosophers think that rigidity cannot be extended from singular terms to general terms. The reason for this is that rigidity appears to become trivial when such terms are considered: natural kind terms come out as rigid, but so do all other general terms, and in particular all descriptive general terms. This paper offers an account of rigidity for natural kind terms which does not trivialize in this way. On this account, natural kind terms are *de jure* obstinately rigid designators and other general terms such as descriptive general terms are not.

A characterisation of rigidity was first offered by Kripke (1971: 144–6), who distinguishes between rigid and nonrigid designators as follows:[1]

> (Kripke Rigidity) A *rigid designator* is a designator that designates the same object in every possible world in which that object exists and does not designate anything in possible worlds in which that object does not exist.
>
> (Kripke Nonrigidity) A *nonrigid designator* is a designator that is not rigid, i.e. a designator that either designates some object in some but not all of the possible worlds in which that object exists or designates different objects in different possible worlds.

There are clear examples of this distinction provided by proper names, such as 'Aristotle', which satisfy (Kripke Rigidity) and definite descriptions, such as 'the pupil of Plato and teacher of Alexander the Great', which satisfy (Kripke Nonrigidity). Of course, there are also many descriptions that satisfy (Kripke Rigidity), such as rigidified or actualized descriptions (e.g. 'the *actual* pupil of Plato and teacher of Alexander the Great') and descriptions that designate necessary existents (e.g. 'the smallest prime'). The latter sort of descriptions belongs to a class of expressions which Kripke calls 'strongly rigid designators' (1980: 48–9): a strongly rigid designator satisfies (Kripke Rigidity) and it designates the same object (e.g. the number two) in every possible because there is no possible world in which that object does not exist—that object is a necessary existent.

Proper names and definite descriptions are singular terms, and one question that has received a lot of attention recently is whether rigidity can be extended to general terms. More precisely the question is whether there are examples of the distinction between (Kripke Rigidity) and (Kripke Non-rigidity) that are provided respectively by natural kind terms (e.g. 'water', 'tiger') and descriptive general terms (e.g. 'transparent liquid of which lakes are composed, and which falls as rain' or 'yellow quadrupedal feline with blackish stripes'). It is natural to expect that while natural kind terms would satisfy (Kripke Rigidity), many descriptive general terms would not, just as while proper names satisfy it, many definite descriptions do not.[2] It is a natural expectation to have because natural kind terms seem to be semantically on a par with proper names, and descriptive general terms seem to be semantically on a par with definite descriptions. For instance, many philosophers think that, like proper names, natural kind terms are both directly referential and nondescriptive and that, like definite descriptions, descriptive general terms are neither. The problem, however, is that (Kripke Rigidity) seems to become trivial when general terms are considered.

Before seeing why, some preliminary remarks about extending (Kripke Rigidity) to natural kind terms are needed.

Firstly, natural kind terms are *general terms*, and some might think that for this reason they should not be treated as referring expressions. Here, I will treat them as such, which is required if they are at all going to be rigid designators: only referring expressions are designators. However, I will briefly consider an alternative proposal following.

Secondly, I will focus on natural kind terms (e.g. 'tiger') rather than on natural kind predicates (e.g. 'is a tiger'), as some do (i.e. Soames 2002). Natural kind terms and predicates are closely related: to every natural kind term there corresponds a natural kind predicate. However, it is more natural to take natural kind terms to be the bearers of fundamental semantic properties such as rigidity or direct reference (on which I will say more later) rather than predicates, which are more complex expressions. I take natural kind terms (or natural kind nouns) to be syntactically and semantically simple expressions whose function is to refer to natural kinds: 'water' is a natural kind term; but descriptions such as 'transparent liquid of which lakes are composed, and which falls as rain' and 'H_2O' are not natural kind terms; they are both syntactically and semantically complex.

Thirdly, for simplicity, I will apply (Kripke Rigidity) to both singular and general terms. This could be objected to on the grounds that (Kripke Rigidity) specifies a relation between a designator and the *object* which it refers to, and it is not obvious that general terms refer to objects. It is indeed not obvious: the fact that an expression refers does not entail that it refers to an object. Intuitively, natural kind terms refer to natural kinds: these may be objects on some construals of natural kinds (e.g. as intensions) but not on others (e.g. as universals).[3] Nothing in what follows turns on whether natural kinds are objects, and so bearing in mind that it might not be a final characterisation when it comes to general terms, I will apply (Kripke Rigidity), as well as other semantic notions defined next, across the board.

Although nothing turns on whether the possible candidates to be the referents of natural kind terms are objects, finding a nontrivial application of (Kripke Rigidity) to natural kind terms puts some requirements on the sorts of properties which these referents should have. In particular, applying (Kripke Rigidity) to natural kind terms requires finding *stable* referents across possible worlds to be the referents of these terms. Here is why. Consider first the simple proposal to construe the natural kinds which natural kind terms refer to as their extensions. On this proposal, natural kind terms will not be rigid because their extensions can be different in different possible worlds. For example the extension of the term 'tiger' is different in different possible worlds because the number of tigers is different in different possible worlds. However, intuitively, 'tiger' refers to the same natural kind in different possible worlds: if there was one tiger less than there actually are, 'tiger' would still refer to the same natural kind. So extensions are not suitable candidates to be the referents of natural kind terms. For an expression to be a rigid designator, its referent needs to stay the same in every possible world in which that referent intuitively exists.

So now consider the proposal that natural kind terms refer to stable entities that do not vary from world to world (*sui generis* natural kinds or properties or intensions or universals or what have you). If we follow this proposal, we are now faced with the problem of *trivialisation*, which I mentioned at the beginning. Natural kind terms will indeed come out as rigid, for instance, 'water' will refer to the same kind in every possible world in which that kind exists. But equally virtually all general terms will come out as rigid, and in particular all descriptive general terms will come out as rigid. The reason for this is that it is just as appropriate to take general descriptive terms to refer to stable referents as it is to take natural kind terms to refer to stable referents. For instance, 'transparent liquid of which lakes are composed and which falls as rain' will be rigid: it will designate the same property (say) in every possible world in which that property exists, namely that of being a transparent liquid of which lakes are composed and which falls as rain.[4]

So rigidity either does not apply (if referents are not stable) or trivializes (if referents are stable) when general terms are considered.

Some have argued that these problems could be avoided by not taking natural kind terms to be referring expressions, a possibility which I have alluded to before. Rather, they are expressions whose function is simply to *apply to objects* or be true of samples in their extensions, and not to refer to natural kinds. For instance, Devitt and Sterelny have proposed the following characterisation of what they call 'rigid application', which is supposed to hold of *all and only* natural kind terms (see Devitt and Sterelny 1999: 85; see also Devitt 2005):

> (Rigid Application) A general term *F* is a rigid applier iff, if *F* applies to an object in the actual world, and that object exists in another possible world, then *F* applies to that object in that world.

(Rigid Application) entails that a rigid applier *F*, if it applies to an object in the actual world, applies to that object in every possible world in which that object exists; that is to say, that object is necessarily *F*. Intuitively given some plausible externalist semantic assumptions, 'water' is a rigid applier because if it applies to samples of water (H_2O) in the actual world, it applies to samples of water in all possible worlds; however, 'transparent liquid of which lakes are composed and which falls as rain' is not: it may apply to samples of water in the actual world and to some other kind in another possible world.

However, this account of rigidity is inadequate. First, (Rigid Application) is not inclusive enough: it leaves out many natural kind terms such as phase sortals (e.g. 'tadpole' or 'caterpillar'), which do not come out as rigid. For instance, 'caterpillar' is not a rigid applier because something that is a caterpillar in the actual world could be a butterfly in some other possible world; that thing is not necessarily a caterpillar.[5] Secondly, (Rigid Application) is too inclusive: it counts as rigid appliers expressions that are not natural kind terms. For instance, it counts noun-phrases such as 'tiger or lemon' as rigid appliers: if 'tiger or lemon' applies to an object in the actual world and that object exists in another possible world, then 'tiger or lemon' applies to that object in that world. Intuitively these are not the sorts of expressions that we should count amongst natural kind terms. So (Rigid Application) is not the correct account of rigidity for natural kind terms.

Because of these difficulties in applying it to general terms, some have concluded that (Kripke Rigidity), and indeed any notion of rigidity, is of no use when it comes to general terms. For instance, Stephen Schwartz has claimed that because natural kind terms and descriptive general terms all come out as (Kripke Rigid), '[r]igidity has lost its exclusivity, like a club of which all are automatically members, and thereby its interest' (Schwartz 2002: 266; see also Soames 2002: Ch. 9). I think that this conclusion is incorrect. In the next sections, I will develop an account of rigidity that is 'exclusive' and that does not trivialize when general terms are considered.

To do so, I will first briefly outline Kaplan's semantic and metasemantic framework of direct reference for proper names. I will then argue that this framework can be used to develop a notion of rigidity that carves at the semantic joints: it applies to proper names and natural kind terms, and not to definite descriptions and descriptive general terms.

2. OBSTINACY AND THE FRAMEWORK OF DIRECT REFERENCE

Many philosophers think that proper names and natural kind terms are directly referential expressions, and that definite descriptions and descriptive general terms are not. A directly referential expression is an expression whose reference is unmediated, i.e. not made via the mediation of any sort of descriptive content. As Kaplan puts it:

> The directly referential term goes directly to its referent, *directly*, in the sense that it does not first pass through the proposition. Whatever rules, procedures, or mechanisms there are that govern the search for the referent, they are irrelevant to the propositional component, to content. When the individual is determined, [it] is loaded into the proposition.[6] (1989b: 569)

Direct and nondirect reference can be defined as follows:

> (Direct Reference) A *directly referential term* is a term that only contributes its referent to the propositions expressed by the sentences in which it occurs.
>
> (Nondirect Reference) A *nondirectly referential term* is a term that refers but does not only contribute its referent to the propositions expressed by the sentences in which it occurs.

For Kaplan, there are in natural language two paradigmatic sorts of directly referential expressions: indexicals and proper names. Both types of expressions only contribute their referents to the propositions expressed by the sentences in which they occur; and for both, the 'rules, procedures, or mechanisms there are that govern the search for the referent' are not part of the propositional content. These rules are for indexicals their *semantic* characters, which determine their referents in different contexts of use. For instance, the semantic character of the indexical 'I' is roughly that 'I' always refers to the agent of the context—where the agent can be different in different contexts. Proper names do not have semantic characters: once fixed, their referents remain the same in all contexts of use. The rules or mechanisms that govern the search for the referent belong to the *metasemantic* facts that are relevant to assigning it a referent in the first place. For instance, once its referent is fixed, the proper name 'Alice' refers to Alice in all contexts of use (See Kaplan, 1989b: 573 ff.).

I will assume that the construal of direct reference which is relevant to natural kind terms is that for proper names: like proper names, natural kind terms do not have semantic characters—their referents are not different in different contexts of use. So although much of what follows would also be true of indexicals, I will confine the scope of my argument to proper names and natural kind terms.

For Kaplan, the metasemantics of direct reference for proper names—that is, the explanation of how they get to be directly referential—involves the notion of *dubbing*. Proper names have the semantic property of direct reference because they are introduced at dubbings. Dubbing can be defined as follows (see Kaplan 1990: 93ff.):

> (Dubbing) A dubbing is the semantic individuation of an expression by assigning it a referent.

Kaplan talks of dubbings as creating words. For instance, the generic name 'Alice' pre-semantically (before it has been used in a particular dubbing) has no meaning whatsoever; if we dub something with it, we create a new name (what he calls a 'common currency name').[7]

In section 4, I will address the questions of what exactly counts as a dubbing and how dubbing might apply to both proper names and natural kind terms. For now, suffice it to say that when it comes to proper names the semantic property of direct reference is entailed by the metasemantic one of being introduced at a dubbing: if a dubbing is the complete semantic individuation of a proper name by assigning it a referent, then all that a name contributes to the propositions expressed by the sentences in which it occurs is the referent that individuates it.

Consider rigidity again. At first, Kaplan took direct reference to be just a way of cashing out (Kripke Rigidity): he thought that the latter was underpinned by the former. However, they come apart. Firstly, not every expression which satisfies (Kripke Rigidity) is directly referential. For instance, rigidified descriptions are not directly referential. Secondly, although every directly referential expression is a rigid designator, it is a different sort of rigid designator than that specified in (Kripke Rigidity). It is an *obstinately rigid designator*. Kaplan defines obstinately rigid and nonobstinately rigid designators as follows (see Kaplan 1989b: 568 ff., and also Salmon 2005: 33–4):

> (Obstinacy) An *obstinately rigid designator* is a designator that designates the same object in all possible worlds regardless of whether it exists at a given world.
>
> (Nonobstinacy) A *nonobstinately rigid designator* is a rigid designator that is not obstinate, i.e. that does not refer to the same object in all possible worlds regardless of whether it exists at a given world.

It is easy to see that (Obstinacy) is a semantic consequence of (Direct Reference): if a directly referential term only contributes its referent to the propositions expressed by the sentences in which it occurs, then, whenever we evaluate these propositions at a world, the referent is already part of the propositions evaluated. So at that world, there automatically is a referent for the term. As Kaplan puts it:

> If the individual is loaded into the proposition (to serve as the propositional component) before the proposition begins its round-the-worlds journey, it is hardly surprising that the proposition manages to find that same individual at all of its stops, even those in which the individual had no prior, native presence. The proposition conducted no search for a native who meets propositional specifications; it simply 'discovered' what it had carried in. (1989b: 569)

Thus (Direct Reference) does not entail (Kripke Rigidity). With (Direct Reference), it is guaranteed in advance of evaluation that a directly referential term has a referent in all possible worlds, something which (Kripke Rigidity) does not guarantee. Consider the sentence 'Aristotle is a philosopher'. If 'Aristotle' is directly referential, Aristotle is part of the proposition expressed by that sentence; if so, unlike with (Kripke Rigidity), when we evaluate that sentence at worlds, we do not look for Aristotle in these worlds, and *then*, if he exists there, say that 'Aristotle' refers to that person in that world and if he does not exist there, say that 'Aristotle' does not refer to anything. With (Obstinacy), if we evaluate 'Aristotle is a philosopher' at a world in which Aristotle does not exist, 'Aristotle' still refers to (the actual) Aristotle. However, the sentence is not true because it is not true that Aristotle is a philosopher in that world (he has no native presence there).[8]

Also, (Direct Reference) does not hold of definite descriptions, which contribute something descriptive to the propositions expressed by the sentences in which they occur: the individuals they refer to are not part of these propositions and at every world at which we evaluate such propositions we have to search for an object that satisfies the description.

Where does this leave us with the issue of definite descriptions and rigidity? Now, although definite descriptions are not directly referential, many such descriptions are obstinately rigid. Of course, some descriptions do not satisfy (Obstinacy): for instance, given that Aristotle does not necessarily exist, 'The pupil of Plato and teacher of Alexander the Great' and its rigidified version 'The actual pupil of Plato and teacher of Alexander the Great' do not satisfy (Obstinacy). The former is not rigid at all, and while the latter satisfies (Kripke Rigidity), it does not satisfy (Obstinacy). This is because 'The actual pupil of Plato and teacher of Alexander the Great' does not refer to anything in possible worlds in which Aristotle does not exist: if nothing in those worlds satisfies the predicate '*x* is the actual pupil of Plato and teacher of Alexander the Great', the definite description is empty.[9] However, descriptions that are strongly rigid designators, such as 'The smallest prime', are obstinate: there will not be worlds in which nothing satisfies the description '*x* is the smallest prime'.

In Kaplan's framework there is an organic connection between (Dubbing), (Direct Reference), and (Obstinacy) for proper names: (Dubbing) entails (Direct Reference), which entails (Obstinacy). But although (Dubbing) and

(Direct Reference) are distinctive of proper names, as opposed to definite descriptions, (Obstinacy) is not, because of the modal status of the referents of strongly rigid definite descriptions. So if we want a notion of obstinacy that is distinctive of proper names, and does not apply to these descriptions, (Obstinacy) is not yet adequate. In the next section, I will consider natural kind terms again, fine tune (Obstinacy), and offer a version of it that applies to proper names and natural kind terms exclusively, and not to descriptive singular or general terms.

3. OBSTINACY *DE JURE* AND OBSTINACY *DE FACTO*

If, like proper names, natural kind terms are directly referential, they are also obstinately rigid. I am going to assume here that both proper names and natural kind terms are directly referential. In section 4, I will discuss this question in greater detail.

Now, recall that our original problem was that (Kripke Rigidity) trivializes when stable referents (such as *sui generis* kinds, intentions, properties, or universals) for natural kind terms are considered. Here, the same sort of problem arises with the notion of rigidity at issue in (Obstinacy). More precisely, the problem comes from the fact that these best candidates to be the stable referents of general terms seem to be *necessary existents*—that is, things that exist in all possible worlds. So (Obstinacy) also trivializes for general terms because of the modal status of the sorts of entities which they refer to: general terms *all* refer to the same things in all possible worlds. In Kripke's terminology, general terms are all strongly rigid.

So (Obstinacy) is not a sort of rigidity that applies to natural kind terms and not to descriptive general terms. This problem with obstinacy is related to that just raised concerning singular terms, where not only proper names but also strongly rigid definite descriptions were obstinately rigid because the latter refer to necessary existents. So both in the case of singular and general terms, the reason why (Obstinacy) is not distinctive of proper names and natural kind terms is that many descriptions satisfy (Obstinacy) because they refer to necessary existents. Although the root of the problem is the same, the problem is more dramatic in the case of general terms because arguably, all descriptive general terms refer to necessary existents and so (Obstinacy) trivializes in this case.

There are several ways in which this problem can be addressed. I will not consider in detail here ways that consist in requiring that the referents of general terms do not necessarily exist. For instance, it could be argued that natural kinds only exist in some possible worlds, namely worlds in which they apply to samples—e.g. worlds in which there are no parcels of water are worlds in which the kind water does not exist. One possible way to construe kinds so understood would be as immanent universals, perhaps in the fashion of Aristotelian moderate realism. On such sorts of proposals, natural kind terms would be obstinately rigid because they are directly referential, and descriptive general terms would not be obstinately rigid because they are not directly referential—many of them would merely satisfy (Kripke Rigidity). In particular, the natural kind term 'water' would be obstinately rigid because it is directly referential; but the descriptive general term 'transparent liquid of which lakes are composed and which falls as rain' would not even be rigid: it would not refer to anything in worlds in which no property satisfies its descriptive content, and it would refer to a different property in worlds in which another property satisfies its descriptive content. For this to be an attractive proposal that works both for general and singular terms, it would also have to hold of the strongly rigid definite descriptions mentioned previously, such as 'The smallest prime'; that is to say, it would also need to be the case that numbers do not necessarily exist.

I do not think that this proposal is worth pursuing. From a semantic standpoint it is not very appealing: it requires us to make substantive assumptions about metaphysical matters, e.g. the metaphysics of kinds or properties or numbers. And it would be more satisfactory to be able to distinguish between the rigidity of natural kind terms and that of descriptive general terms not only on metaphysical, but also on semantic grounds.

Rather, I suggest that we look for a construal of (Obstinacy) that avoids trivialisation by appealing to Kripke's distinction between *de jure* and *de facto* rigid designation, which he applies to his own characterisation of rigidity, (Kripke Rigidity). According to Kripke, a designator is rigid *de jure* if it is rigid as a matter of *stipulation* and it is rigid *de facto* if it happens to be rigid. As he puts it, the distinction is between:

> '*de jure*' rigidity, where the reference of a designator is *stipulated* to be a single object . . . and mere '*de facto*' rigidity, where a description 'the x such that Fx' happens to use a predicate 'F' that in each possible world is true of one and the same unique object. (Kripke 1980: 21 n.21)

For instance, the description 'The smallest prime' satisfies (Kripke Rigidity) *de facto* because it merely happens to use the predicate 'is a smallest prime', which in each possible world is true of the number two. By contrast, although the description 'The number Alice is thinking about now' happens to refer to the number two, it is not *de facto* rigid because there are worlds in which 'is a number which Alice is thinking about now' refers to another number. Indeed, that latter description does not satisfy (Kripke Rigidity). The way Kripke states it, *de facto* rigidity only holds of strongly rigid designators—rigid designators of necessary existents—because it demands that the predicate in the description is true of one and the same unique object *in each possible world*. However, it is natural to think that descriptions that satisfy (Kripke Rigidity) but are not strongly rigid could still be rigid *de facto*: intuitively 'The actual pupil of Plato' satisfies (Kripke Rigidity) *de* *facto*, although it is not strongly rigid: it happens to use a predicate 'is an actual teacher of Plato' that is true of one and the same unique object in every possible world *in which that object exists*. I.e. since there are worlds in which that object does not exist, that description is not strongly rigid.

I will assume here that we can apply *de facto* rigidity to descriptions that are rigid but not strongly rigid. However, note that for my purposes nothing turns on this because the descriptions that are troublesome for (Obstinacy) as the distinctive notion of rigidity for proper names and natural kind terms are all *de facto* strongly rigid. More precisely, only definite descriptions and descriptive general terms that satisfy (Kripke Rigidity) *de facto* and are strongly rigid satisfy (Obstinacy). Definite descriptions and descriptive general terms that satisfy (Kripke Rigidity) *de facto* but are not strongly rigid designators do not satisfy (Obstinacy).

Thus, two things conspire in making (Obstinacy) not distinctive of proper names and natural kind terms: the fact that certain descriptions use certain predicates, which given their meanings happen to apply to the same thing in all possible worlds, and the metaphysical status of the things referred to by the descriptions that contain those predicates.

Let us now apply the *de jure–de facto* distinction to (Obstinacy), which is the construal of rigidity required by (Direct Reference) and so intuitively applies to proper names and natural kind terms. We get the following contrast between two ways of being obstinately rigid:

> (*De Jure* Obstinacy) A *de jure obstinately rigid designator* is a designator that designates the same object in all possible worlds as a matter of stipulation.
>
> (*De Facto* Obstinacy) A *de facto obstinately rigid designator* is a designator that designates the same object in all possible worlds, but not as a matter of stipulation.

My suggestion is now that (*De Jure* Obstinacy) is the distinctive notion of rigidity we are looking for: both proper names and natural kind terms satisfy (*De Jure* Obstinacy), and neither definite descriptions nor descriptive general terms do—although they may satisfy (*De Facto* Obstinacy). In particular (*De Jure* Obstinacy) does not trivialize when general terms are considered. Also, whether a designator satisfies (*De Jure* Obstinacy) has nothing to do with the metaphysical status of its referent, i.e. with whether it refers to a necessary existent. It has to do with the fact that it is stipulated to be obstinately rigid.

To make this suggestion precise, the notion of a stipulation needs to be made precise, which turns out to be no straightforward matter.

The contrast between *de jure* and *de facto* rigidity is formulated in terms of stipulation and lack of stipulation as to whether a term is rigid. As it is, this is unsatisfactory because there is a sense in which whether an expression is rigid is always a matter of stipulation—or at least, there is always a sense in which a stipulation comes into play. And this is the case whether or not that expression is directly referential. Thus consider, for instance, the case of rigidified descriptions such as 'the actual pupil of Plato and teacher of Alexander the Great'. There is a clear sense in which such descriptions are rigid as a matter of stipulation—because the adjective 'actual' is defined so as to always refer back to the actual world. So there is a sense in which it does not seem to be the case that such descriptions are rigid because they contain predicates (e.g. 'is an actual pupil of Plato and teacher of Alexander') that *happen* (i.e. is not stipulated) to apply to the same objects in every possible world in which those objects exist. Such descriptions have been *rigidified*, i.e. stipulated to be rigid.

Now, again, descriptions such as 'the actual pupil of Plato and teacher of Alexander the Great' do not immediately concern us here because they satisfy (Kripke Rigidity) but not (Obstinacy). However, consider again Kripke's example of a *de facto* strongly rigid designator, such as the description 'the smallest prime', or better the rigidified description 'the actual smallest prime'. It is not obvious that these descriptions do not satisfy (*De Jure* Obstinacy). So the distinction between *de jure* and *de facto* demands further analysis.

One possible articulation of the distinction between *de jure* and *de facto* obstinate rigidity goes by appealing to the notion of contingency, which is invited by that of a description happening to use a predicate that is true of one and the same unique object in each possible world. On this proposal, the contrast would be between expressions that are contingently rigid (*de facto*) and expressions that are not contingently rigid (*de jure*)—where the former are expressions that, in some sense, could have failed to be rigid. On this proposal, 'the smallest prime' would be a description that could have failed to be obstinately rigid because the predicate 'x is a smallest prime' could have failed to be true of the same unique object in each possible world.

There are two ways of making this proposal. The first would yield that 'the smallest prime' is contingently obstinately rigid because it could have failed to refer to a necessary existent, namely the number two. But this cannot be made sense of: it is false to say that the 'the smallest prime' could have failed to refer to the number two; in particular it is false to say that 'x is a smallest prime' could have failed to apply to the number two. Given this description's (and predicate's) meaning, it just does not seem that it could have failed to refer to that number: it would have needed to have another meaning for it to do so. Here it would not help to individuate expressions syntactically, and say that rigidity is a property of syntactic type, rather than semantically in terms of their meanings: 'the smallest prime' would indeed only be contingently rigid, but then all rigid expressions would only be contingently rigid. In particular all expressions that we hoped would be rigid *de jure*, such as proper names and natural kind terms.

The second way has it that 'the smallest prime' is contingently obstinately rigid because the number two might not have existed necessarily. On this proposal, the description 'the smallest prime' would be contingently rigid because if the number two had not been a necessary existent, the description would not have been rigid. The problem here is again that this option forces us to make decisions about the metaphysical status of the number two. Also, more generally, it requires us to deny that what is necessary is necessarily necessary, i.e. to reject S4, something that perhaps we should not do (at least not for this reason).

So we cannot articulate the contrast between *de jure* and *de dicto* obstinacy in terms of whether it is contingent that an expression satisfies (Obstinacy).

Consider yet another way of articulating the distinction between *de jure* and *de facto* rigid designators. It is sometimes suggested that this distinction is just that between *nondescriptive* and *descriptive* designators.[10] However, the suggestion of assimilating *de jure* rigidity to nondescriptive designation and *de facto* rigidity to descriptive designation is unsatisfactory; for although it might in fact be true that all expressions that are *de jure* rigid are nondescriptive and that all expressions that are *de facto* rigid are descriptive, nevertheless these two distinctions—*de jure*–*de facto* rigidity and nondescriptive–descriptive designation—are intuitively different. The former concerns the way an expression is made to refer to its referent and the latter concerns whether an expression is semantically complex. If these are distinct semantic distinctions (although perhaps coextensive), we need to find something else to explain the distinction between *de jure* and *de facto* rigidity, as well as the fact that it might coincide with that between nondescriptive and descriptive designation. It is true that, given the way *de facto* rigidity is defined by Kripke, only descriptive expressions are *de facto* obstinate, for whether an expression is *de facto* obstinate is a matter of it using a predicate that happens to apply to the same thing in all possible worlds. If an expression uses a predicate, it is descriptive. However, the way the distinction between *de jure* and *de facto* is stated does not *preclude* there being descriptions that are *de jure* obstinate. So we still need an explanation of why all *de jure* obstinate expressions might be nondescriptive.

I now want to argue that the contrast between *de jure* and *de facto* obstinate rigid designation should be explained in terms of the metasemantics of direct reference. I think that it is natural to turn to direct reference for an explanation of the distinction between *de jure* and *de facto* rigidity: direct reference semantically entails obstinate rigid designation, and it is natural to think that this entailment holds as a matter of stipulation. Now, as we have seen, (Direct Reference) applies to proper names because they are introduced at dubbings. I suggest that it is (Dubbing) that is the relevant metasemantic explanation for why an expression has the semantic property of direct reference. Also, the notion of a stipulation is intuitively a metasemantic notion, which concerns our intention/decision to assign a given semantic property to an expression. So it is appropriate to turn to dubbing as the relevant metasemantic notion to explain *de jure* obstinacy. Again, a dubbing is the semantic individuation of an expression by assigning it a referent. And so an expression introduced as a dubbing is directly referential: it only contributes its referent to the propositions expressed by the sentences in which it occurs. And that means that that expression is obstinately rigid—more precisely, it is *de jure* obstinately rigid because it is the very way in which it has been individuated that makes it obstinately rigid.

This metasemantic explanation in terms of dubbing also gives us a good contrast between expressions that satisfy (*De Jure* Obstinacy) and those that merely satisfy (*De Facto* Obstinacy). As I said, I take it that proper names and natural kind terms satisfy the former and definite descriptions and descriptive general terms do not. Definite descriptions and descriptive general terms are not introduced at dubbings: they are not individuated by assigning them referents. So they are not stipulated to be obstinate. Such expressions will merely be *de facto* obstinately rigid designators, if they refer to things that exist in all possible worlds. Further, the metasemantic explanation of (*De Jure* Obstinacy) in terms of dubbing also explains why expressions that are *de jure* obstinate are not descriptive: for descriptive expressions are not individuated by assigning them a referent. So *de jure* obstinacy and nondescriptivity indeed coincide, but the explanation for this is given from above—in terms of the fact that they are semantic consequences of the same metasemantic story.

So the final proposal concerning rigidity for proper names and natural kind terms is this: it is (*De Jure* Obstinacy) that is the distinctive, nontrivial notion of rigidity. Natural kind terms and proper names are *de jure* obstinately rigid; definite descriptions and descriptive general terms are not. To satisfy (*De Jure* Obstinacy) such terms have to satisfy (Direct Reference) and (Dubbing), because the stipulation alluded to in (*De Jure* Obstinacy) is the dubbing. Ultimately it is the metasemantics of direct reference that enables us to single out proper names and natural kind terms as *de jure* obstinately rigid designators. And in the next section, I say more about how the metasemantic notion of dubbing should be taken to work.

In this debate concerning whether rigidity trivializes when general terms are considered, it is generally assumed that in order to draw a contrast between two sorts of expressions in terms of (some construal of) rigidity that carves at the semantic joints, we ought to say that the one sort is rigid and that the other is *not at all* rigid. The current proposal makes no such assumption, and given the plethora of notions of rigidity that are available, there is no reason to expect that such an assumption should hold. According to my proposal we carve at the semantic joints by distinguishing between different types of rigidity. Both proper names and natural kind terms are distinctive in satisfying (*De Jure* Obstinacy), descriptive expressions that refer to necessary existents satisfy (*De Facto* Obstinacy), and remaining rigidified descriptions satisfy (Kripke Rigidity). And so it may be the case that rigidity is like a club which is not very interesting because too many get automatic membership. However, if I am right, there is still a very exclusive VIP area at the back of the club, where one gets in only by stipulation.[11]

4. DUBBING AND NATURAL KINDS TERMS

I have claimed that the distinctive notion of rigidity for proper names and natural kind terms is that of *de jure* obstinate rigidity. And I have argued that for them to be rigid in this way, they would have to be introduced at dubbings. In this last section I briefly address some possible worries concerning the idea that natural kind terms might be introduced at dubbings. Here I assume that the claim that proper names are introduced at dubbings to be comparatively unproblematic, and so I do not offer a defence of the claim that they are. I merely argue that saying that natural kind terms are introduced at dubbings is no more problematic than saying that proper names are introduced at dubbings. A full defence that proper names and natural kind terms are introduced at dubbings would require a separate paper.

As I said before, for Kaplan a dubbing is the semantic individuation of a term by assigning it a referent. One key condition on the successful introduction of a term at a dubbing is that there is a unique referent for that term. The other is that the dubber intends to introduce a new term rather than to follow an already established use of a term (see Kaplan 1989a: 560). I label these two conditions '(Creativity)' and '(Unicity)':

(Creativity) The dubber has *creative linguistic intentions*.
(Unicity) There is a *unique referent* to the term introduced.

(Unicity) entails that a term that does not uniquely refer is not the result of a successful dubbing. In particular, empty terms are not the result of successful dubbings: for instance, terms that—perhaps unbeknownst to the would-be dubber—turn out to be empty at the dubbing or terms that are introduced in fictional contexts.[12] That means that (Unicity) can trump (Creativity); the creative linguistic intentions of the dubber are *defeasible*: she may be wrong—if there is no unique candidate to be the referent, she may intend to introduce a term and fail to do so. Saying that creative linguistic intentions are defeasible implies that those intentions do not carry a huge amount of semantic weight. As a corollary, that implies that it is not always transparent to a speaker which semantic properties an expression has, in particular whether that expression is directly referential or is a rigid designator (and what sort). This might be hidden from view at the time of the dubbing.

These conditions on dubbing give us a good *prima facie* contrast between proper names and definite descriptions. For one thing, (Creativity) has no clear application to the latter, because definite descriptions are typically made up of expressions that are already in the language. So there is no real introduction or creation of such a description, but rather the putting together of expressions that have already been assigned a meaning. Moreover, (Unicity) is not a condition on the use of a definite description, whose meaning is not tied to there being a unique referent: definite descriptions can be empty or can pick out several objects.

The question now is whether these conditions on dubbing give us a good contrast between natural kind terms and general descriptive terms. I now consider three objections to the idea that natural kind terms are introduced at dubbings, and argue that they are not successful.

Firstly, it could be argued that natural kind terms are typically not introduced at dubbings for the reason that a dubbing is a simple unique act of baptism, and natural kind terms are generally not introduced in the language by some single such act. They typically emerge in the language as the result of a more gradual process, not as the result of a single intentional act but of many.

That seems right. However, note that the same holds of many names—many names are gradually introduced in the language. For instance, we can think of the gradual process by which a nickname sticks (Evans's example in his 1973). So it is not clear that we have a significant difference here between proper names and natural kind terms. In both cases, talk of 'dubbing', if it suggests a single act of baptism, is a convenient fiction, and it would ultimately have to be understood in a way that allows for directly referential expressions to be gradually introduced in the language.[13]

One helpful way to understand the notion of a dubbing is as follows: saying that a term has been introduced at a dubbing need not be making an actual historical claim about its actual mode of introduction; e.g. it need not be making the claim that there was actually a single act of baptism during which that term was introduced. For many terms, we just do not know the precise context or mode of their introduction into the language, and probably such an introduction was somewhat messy. For such terms, the idea of them being introduced at dubbings could rather be taken as follows: given their basic semantic properties (e.g. direct reference, *de jure* obstinacy), these terms are such that they *could* have been introduced at a single act of baptism, i.e. at a dubbing. For instance, given the basic semantic properties of 'tiger', this natural kind term could have been introduced at a single act of baptism. In particular, 'tiger' could have been introduced by just pointing at a tiger and ostensively define the term 'tiger' by saying

'Here is a tiger' or 'A tiger is anything like this'. If a term could intuitively have been introduced in this way, we have a good explanation of why it is not descriptive, but directly referential. So dubbing so understood gives us a good explanation of why certain terms have the basic semantic properties that the do. It gives us a good theoretical reconstruction or metasemantic explanation of what is semantically distinctive about certain terms, such as natural kind terms.

Secondly, one might argue that natural kind terms are not introduced at dubbings because although we can use proper names to arbitrarily name anything at will, we cannot do this with natural kind terms. Intuitively, the latter cannot just be created at will; they have to refer to natural kinds, and not just anything is a natural kind: stipulations of referents for proper names come cheap, but not so for natural kind terms.

However, note first that Kaplan's account of proper names is in any case inhospitable to the idea that we can create proper names at will: (Unicity) would require that there is a unique object that is dubbed using a proper name, and so creative intentions are defeasible in the case of proper names. What is true, though, is that the class of thing that can be successfully dubbed with a name is bigger than the class of things that can be successfully dubbed with a natural kind term—natural kinds are harder to come by and so natural kind terms are harder to come by. But that does not seem to be a semantically significant fact. For instance, Danish proper names are harder to come by than English proper names, but that does not point to any significant semantic difference between them.

Thirdly, a related objection could be that it is somehow more difficult to introduce natural kind terms than it is to introduce proper names because the former refer to complex things, things with a complex (molecular, biological, or what have you) structure, while the latter do not. One way of making this objection is by saying that the *intentions* involved in introducing a natural kind term are more complex than those involved in introducing a proper name because one intends to dub something with a complex structure.

But this is rather contentious. Consider, for instance, the natural kind term 'water'. Putnam (1975) has convincingly argued that 'water' did refer to the kind water at the time at which it was introduced in the language, although that time was well before people had any knowledge of chemistry, or even any idea of the sort of complexity that natural kind might involve. So it does not seem that in the case of natural kind terms the intentions of the dubber(s) ought to be more complex than in the case of proper names.

Another way of making this objection from complexity is to say that dubbing a natural kind is more complex not because the intentions of the dubber(s) have to be more complex but merely because the things dubbed are more complex. Their complexity is such that, typically, it is empirical, scientific, investigation that reveals whether the dubbing has been successful—whether a unique genuine natural kind has been referred to.

However, and this is relevant to the previous paragraph, we can of course dub complex things using proper names. In particular, we can dub natural kinds using proper names, where perhaps only empirical investigation could tell whether the dubbing is successful. And we might also want to dub objects with a complex structure using proper names. For instance, we might want to dub with the name 'Alice' (say) something which we think is a fertilized egg (i.e. a human animal); but we might fail to do so because there is no fertilized egg to be dubbed but only a bunch of cells that do not (yet) form a single organism. Only empirical investigation will reveal that the introduction of the name 'Alice' was unsuccessful.

So there does not seem to be significant differences between proper names and natural kind terms with respect to the way (Dubbing) applies to them. If so, the claim that natural kind terms, just like proper names, are introduced at dubbings is plausible.

The notion of dubbing as explained here also further enables us to draw a satisfactory contrast between natural kind terms and descriptive general terms. Like definite descriptions, descriptive general terms are typically made up of expressions already present in the language, so they are not as such introduced in the language. So (Creativity) typically does not apply. But also (Unicity) is not a condition on the use of descriptive general terms: descriptive general terms can be empty or pick out several kinds, or things. A successful use of a descriptive general term such as 'transparent liquid of which lakes are composed and which falls as rain' does not require the existence of a unique kind.

Given the connection between (Dubbing), (Direct Reference), and (*De Jure* Obstinacy), if it can plausibly be said that natural kind terms indeed satisfy (Dubbing) and descriptive general terms do not, that means that the former satisfy (*De Jure* Obstinacy) and the latter do not. And so, again, we have got hold of a notion of rigidity that carves at the semantic joints.

5. CONCLUDING REMARKS

In this paper, I have argued that proper names and natural kind terms are distinctive in being *de jure* obstinately rigid designators. What explains that they have this semantic property is that they are directly referential terms introduced at dubbings. Dubbing is the metasemantic parameter that ultimately explains the distinctive status of proper names and natural kind terms as *de jure* obstinately rigid designators. Some descriptive general terms and definite descriptions are rigid (in some sense or other of rigidity), but given that they are not introduced at dubbings, they are not *de jure* obstinately rigid.[14]

APPENDIX

The picture of rigidity that we get if we adopt the considerations put forward in this paper are summarized in the following table. Note that many of the expressions that appear in the squares are just standard examples for illustration. For instance, I have taken 'water' to refer to a necessary existent and put it in the square labelled 'I' (*de jure* strongly obstinately rigid designator), but some might rather put it in the square labelled 'III' (*de jure* non-strongly obstinately rigid designator). Also, not all the expressions that figure in this table have been discussed in the paper (e.g. demonstratives and indexicals) but I included them for completeness. The aim here is just to give a general idea of how a classification of different types of rigid designators might work.

Table 2.1 ?

	Obstinately rigid designator	*Rigid designator, nonobstinate*
***De jure* + strongly** (originates at a dubbing + refers to a necessary existent)	'Water' 'Tiger' '2' **I**	——— **II**
***De jure* + non-strongly** (originates at a dubbing + does not refer to a necessary existent)	'Aristotle' 'I' 'That' 'dthat' **III**	——— **IV**
***De facto* + strongly** (does not originate at a dubbing + refers to a necessary existent)	'The smallest prime' 'The actual smallest prime' 'Transparent liquid of which lakes are composed, and which falls as rain' 'H2O' **V**	——— **VI**
***De facto* + non-strongly** (does not originate at a dubbing + does not refer to a necessary existent)	——— **VII**	'The actual pupil of Plato and teacher of Alexander the Great' **VIII**

NOTES

1. What I call here 'Kripke Rigidity', Salmon calls 'persistent rigidity' (See 2005: 34). Note here the difference between Kripke's first characterisation of rigidity in his 1971: 146, i.e. my (Kripke Rigidity), and that which he gives in *Naming and Necessity*: '[A] designator rigidly designates a certain object if it designates that object wherever [in every possible world in which] the object exists' (1980: 49). Unlike the first one, this second characterisation is neutral on whether a rigid designator still designates the same object with respect to possible worlds in which that object does not exist. That is to say, it is neutral on whether rigid designators are persistently/Kripke rigid or obstinately rigid. Obstinate rigidity will be introduced in section 2. It should be stressed that Kripke distanced himself from the characterisation given in his 1971, which is the view of rigidity usually attributed to him, in favour of the more neutral one he put forward in his 1980. See Kaplan 1989b, n. 8 and my n. 8 for further discussion.
2. See *inter alia* Kripke 1980, Donnellan 1983, Schwartz 2002, Soames 2002, Salmon 2003, 2005, and Martí 2004 for discussions of rigidity for natural kind terms. I leave aside the issue of how terms for artefacts (e.g. 'clock') or cultural properties (e.g. 'bachelor') should be understood: whether they should be assimilated to natural kind terms or to descriptive general terms.
3. Some natural kind terms are mass nouns (e.g. 'water') and some are count nouns (e.g. 'tiger'). It is a debated question how different syntactically and semantically the two sorts of terms really are (see Koslicki 1999 for discussion). For instance, on some accounts 'water' is a singular term that refers to an object; on others it is a general term that does not refer to an object. I shall treat both mass nouns and count nouns as general terms. Nothing substantive hangs on doing so here; for if mass nouns are singular terms, the conception of rigidity for natural kind terms and proper names developed in this paper still apply to them.
4. Several attempts have been made (e.g. Martí 2004 and LaPorte 2006) to try to create an asymmetry between natural kind terms and descriptive general terms by distinguishing between different properties that the latter might refer to (or express), some rigidly, others nonrigidly. I do not review these proposals here.
5. It could be denied that phase sortals are natural kind terms (Devitt 2005); but this seems *ad hoc*. It could also be claimed that rather than applying to (enduring) objects, natural kind terms apply to temporal parts: in such a case 'caterpillar' would apply rigidly to all and only the temporal parts that are in its extension; any such temporal part would necessarily be a caterpillar. I do not discuss this option further, which seems to be an overreaction to the problem of applying rigidity to natural kind terms.
6. Kaplan is working with a Russellian construal of propositions; other construals of propositions would be possible as long as there is a clear sense in which objects can be components of these propositions.
7. See Kaplan 1989a: 560–1, 1973: 500 ff., and 1990, where he argues that the relation between generic and common currency names is not that between type and token: type and token are type and token of the same expression syntactically or syntactico-semantically individuated; but the relation between generic and common currency names is merely orthographic. We do not have to choose here between these ways of understanding this relation, which can both make sense of the idea that proper names are directly referential.
8. Kaplan initially thought that direct reference underpinned Kripke's notion of rigidity because he thought that Kripke meant by it what he means by

'obstinate rigidity' (see Kaplan 1989b: 569–71). Not so. In a letter that Kaplan partly reproduces (n. 8), Kripke gives the following definition of rigid designation: 'a designator *d* of an object *x* is rigid, if it designates *x* with respect to all possible worlds where *x* exists, and *never designates an object other than x with respect to any possible world*' (Kripke's italics). Like what I have called 'Kripke Rigidity' in section 1, this definition does not require it to be settled in advance of evaluation what a designator refers to at worlds at which there is no candidate to be the name's referent. However, unlike with (Kripke Rigidity), an interpretation of this definition in terms of obstinacy is not ruled out.

9. Here, I assume that the domain of the variable with respect to a possible world is restricted to the individuals in that world.
10. Salmon (2005: 35) takes the distinction in this way. And in many ways it is a natural way to take what Kripke says about it (see again his 1980: 21, n. 21). See also Stanley 1997: 537.
11. In the appendix, a table summarizes the different notions of rigidity and which sorts of expressions they respectively apply to.
12. Kaplan wants to rule out expressions that are introduced in fictional, imaginary, hallucinatory, or otherwise defective contexts as contexts in which an individual is dubbed. In his 1968: 383 he also excludes as successful dubbings of merely possible objects such as if we dubbed 'Newman 1' the first child born in the twenty-second century. (However, this is counted as a successful dubbing in his 1973: n. 7.) This sort of case raises the question of the extent to which the dubber has to be acquainted (whether demonstratively or descriptively) with the dubbee in order for the dubbing to count as a dubbing *of* the dubbee. See Kaplan 1968 and 1973 for discussion.
13. Many philosophers talk of reference-fixing or ostensive definitions in connection with natural kind terms–these too suggest a single act of baptism, and these too are convenient fictions. Note in passing that dubbing and reference-fixing are different. Unlike dubbing, reference-fixing need not be understood as the semantic individuation of a term, but merely as the fixation of the referent of that term. So the latter is in principle compatible with the idea that we have a meaningful natural kind term in advance of assigning it a referent, and in principle it allows for there to be meaningful empty natural kind terms (terms that are natural kind terms and that are empty)—it thus does not entail direct reference or obstinate rigidity.
14. Thanks to Brian Ball, Michael Blome-Tillmann, Geoffrey Ferrari, Gail Leckie, Hemdat Lerman, James Morauta, Bruno Whittle, and Timothy Williamson for very useful discussions on a draft of this paper. Thanks also to Alexander Bird, Jane Friedman, and audiences at the Nature and Its Classification conference in Birmingham, at the SEFA 5 in Barcelona, and at the 2007 Joint Session of the Aristotelian Society and Mind Association in Bristol. Special thanks to Helen Beebee and Nigel Leary for very helpful detailed comments on the penultimate draft of this paper.

REFERENCES

Devitt, M. (2005) 'Rigid application', *Philosophical Studies*, 125: 139–65.
Devitt, M. and Sterelny, K. (1999) *Language and Reality: An Introduction to the Philosophy of Language*, 2nd edn. London: Basil Blackwell.
Donnellan, K. (1983) 'Kripke and Putnam on natural kind terms', in C. Ginet and S. Shoemaker (eds), *Knowledge and Mind: Philosophical Essays*, Oxford: Oxford University Press, 84–104.

Evans, G. (1973) 'The causal theory of names', in his *Collected Papers*, Oxford: Oxford University Press, 1–24.

Kaplan, D. (1968) 'Quantifying in', in P. Martinich (ed.), *The Philosophy of Language*, Oxford: Oxford University Press, 371–91.

———. (1973) 'Bob and Carol and Ted and Alice', in J. Hintikka (ed.), *Approaches to Natural Language*, Dordrecht, Netherlands: Kluwer, Synthèse Library, 490–518.

———. (1989a) 'Demonstratives', in J. Almog and J. Perry (eds), *Themes from Kaplan*, Oxford: Oxford University Press, 481–563.

———. (1989b). 'Afterthoughts', in J. Almog and J. Perry (eds), *Themes from Kaplan*, Oxford: Oxford University Press, 565–614.

———. (1990) 'Words', *Proceedings of the Aristotelian Society*, Supp. vol. 44: 93–119.

Koslicki, K. (1999) 'The semantics of mass predicates', *Noûs*, 33: 46–91.

Kripke, S. (1971) 'Identity and Necessity', in M. K. Munitz, (ed.), *Identity and Individuation*, New York: New York University Press, 135–64.

———. (1980) *Naming and Necessity*. Oxford: Blackwell.

LaPorte, J. (2006) 'Rigid designators for properties', *Philosophical Studies*, 130: 321–36.

Martí, G. (2004) 'Rigidity and general terms', *Proceedings of the Aristotelian Society*, 104: 129–46.

Putnam, H. (1975) 'The meaning of "meaning" ', in his *Mind, Language and Reality: Philosophical Papers* vol. 2, Cambridge: Cambridge University Press, 215–71.

Salmon, N. (2003) 'Naming, necessity, and beyond', *Mind*, 112: 475–92.

———. (2005) *Reference and Essence*, 2nd edn. New York: Prometheus Books.

Schwartz, S. (2002) 'Kinds, general terms, and rigidity', *Philosophical Studies*, 109: 265–77.

Soames, S. (2002) *Beyond Rigidity: The Unfinished Agenda of Naming and Necessity*. Oxford: Oxford University Press.

Stanley, J. (1997) 'Names and rigid designation', in B. Hale and C. Wright (eds), *A Companion to the Philosophy of Language*, Oxford: Blackwell, 555–85.

3 General Terms as Designators
A Defence of The View

Genoveva Martí and José Martínez-Fernández

1. INTRODUCTION

According to a traditional and pervasive view in semantics, general terms such as 'gold', 'tiger', 'philosopher', 'computer', or 'Mary's favourite colour' designate universals: abstract entities such as substance and artifact kinds, colours, and species. This view (from now on *The View*) may be metaphysically unpalatable, especially for those of a nominalist persuasion, since it goes hand and hand with what may be seen as an overpopulated ontology. However, The View has, as of late, come into disrepute, not for metaphysical reasons, but for purely semantic reasons: it is argued that The View does not allow for a significant characterization of the notion of *rigidity* for general terms. The View appears to suggest a very natural extension of the notion of rigidity that Kripke defined only for singular terms. Singular terms are rigid just in case they designate the same individual with respect to all indices of evaluation, or possible worlds. Since The View treats general terms as designators, it is natural to define the rigidity of general terms in a similar way: a general term is rigid just in case it designates the same universal with respect to all indices of evaluation. In principle this characterization seems to accord with our initial judgments about rigidity: 'blue' is rigid because it designates the same colour with respect to all possible worlds, whereas 'Mary's favourite colour' is surely not rigid: had Mary's taste been different, 'Mary's favourite colour' would have designated a different colour.

But, it has been pointed out, things don't quite work out as expected. The View's characterization of rigidity has three major problems. First of all, we are told, The View trivializes the notion of rigidity. It is ultimately impossible for The View to distinguish rigid from non-rigid readings of general terms: when we consider expressions such as 'the colour of the sky', it makes no difference, from a semantic point of view, whether we interpret the term as non-rigid, i.e. as designating different colours in different possible worlds, or as rigid, i.e. as designating in every possible world the universal that things instantiate in a world w when they have the property of being the same colour as the sky in w. If 'blue' rigidly designates a universal, the colour blue, then 'Mary's favourite colour' or 'the colour of the sky' can also be interpreted as rigidly designating a universal. Any sentence of the form *t is G* (or *t is a G*) is true just in case *t* instantiates *G-ness*, so any general term *G*, arguably, designates *G-ness* and it does so rigidly. On this approach all general terms may as well be considered rigid.[1]

The trivialization problem is, no doubt, the most serious challenge faced by The View, but even if there was a solution to the trivialization problem, The View faces other challenges. It is argued that an adequate definition of the notion of rigidity for general terms must satisfy three desiderata: (i) it must be a natural extension of the notion of rigidity for singular terms; (ii) it must classify natural kind terms and terms for natural phenomena as rigid and classify many other general terms as non-rigid; and (iii) it must account for the necessity of true theoretical identifications involving rigid terms.[2] Even if there was an adequate response to the charge of trivialization, The View, allegedly, would fall short of satisfying desiderata (ii) and (iii).

The second alleged problem is that The View overgeneralizes rigidity. Natural kind terms such as 'gold' and 'water' do come out rigid according to The View, but so do non-natural kind terms, such as 'computer', 'bachelor', and 'philosopher', and those, according the second desideratum, should not be classified as rigid.[3]

Thirdly and finally, it is argued that The View does not satisfy the third, and crucial, desideratum, for it does not account for the necessity of true theoretical identifications. Statements of the form t_1 *is* t_2 are necessary if true whenever t_1 and t_2 are rigid singular terms. But if t_1 and t_2 are general terms, their rigidity does not guarantee the necessity of the true theoretical identifications expressed by sentences containing them. Sentences such as 'cats are mammals' or 'water is H_2O' are sentences whose forms are respectively (a) $\forall x(Gx \rightarrow Hx)$ and (b)$\forall x(Gx \leftrightarrow Hx)$. Even if *G* and *H* rigidly designate universals, (a) and (b) may well be contingently true, since *G* and *H* may designate universals that happen, but just happen, to be instantiated by the same objects or samples.

Those are three important semantic challenges to The View. Even though Kripke never fully characterized what rigidity for general terms is supposed to consist in, there is consensus among semanticists that the notion of rigidity is applicable to general terms, and there is a substantial level of agreement, at least as regards a good number of paradigm cases, about whether certain terms are rigid or non-rigid. Thus, a view that falls prey to those three fundamental objections should perhaps be abandoned.

We think that The View's take on rigidity can meet the challenges. Our purpose in this paper is to address and respond to the three fundamental objections.

2. THE TRIVIALIZATION OF RIGIDITY

Joseph LaPorte has argued that The View does not trivialise the notion of rigidity:

> My preferred account does not trivialize rigidity. It is simply not the case that every kind designator rigidly designates its kind . . . 'soda pop', 'soda', and 'pop' all rigidly designate the soda pop kind; but 'the beverage my uncle requests at Super Bowl parties' only accidentally designates the kind . . . (LaPorte 2000: 296)
>
> 'Soda = the beverage my uncle requests at Super Bowl parties' is true but not necessarily true, since the second designator is not rigid. (LaPorte 2000: 299).

In the most natural interpretation, the sentence 'Soda is the beverage my uncle requests at Super Bowl parties' is only contingently true and so 'the beverage my uncle requests at Super Bowl parties' is a non-rigid designator that designates different kinds of beverages in different possible worlds.

Examples like this one show that there is no trivialization problem when general terms are used as terms,[4] i.e. when the sentences at stake are naturally interpreted as making claims *about* universals (about kinds, colours, substances, or, as in this case, types of beverages). Similarly, 'blue is the colour of the sky' is contingently true, and this shows that 'the colour of the sky' designates different colours in different possible worlds.

But the problem of trivialization persists when general terms are used as predicates. Consider, for instance, the sentence

(1) Ann's dress is the colour of the sky

Here 'the colour of the sky' is used as a predicate that is attributed to Ann's dress. In the non-rigid reading, and according to The View, 'the colour of the sky' designates the colour blue in the actual world and perhaps the colour red in a different world *w*. Sentence (1) will be true at a world if, and only if, Ann's dress instantiates the universal denoted by 'the colour of the sky' at that world. So (1) is true in the actual world if Ann's dress is blue, and it is true in *w* if Ann's dress is red.

But 'the colour of the sky' also has a rigid reading on which it denotes, at all possible worlds, the same universal: the universal that a thing instantiates when it has the property of being the same colour as the sky.[5] In intensional semantics, this universal would be represented as a function that assigns to the actual world the set of blue things in the actual world, to *w* the set of things red in *w*, etc. So (1) will be true at a world if, and only if, Ann's dress exemplifies the property of being the same colour as the sky in that world. In other words, (1) will be assigned true in the actual world if Ann's dress is the same colour as the sky in the actual world and it will be true with respect to *w* if Ann's dress is the same colour as the sky in *w*.[6]

It is easy to see why the trivialization problem persists: the condition associated with the rigid reading of 'the colour of the sky' assigns to (1) the same truth values as the condition associated with the non-rigid reading: (1) will be true in the actual world if Ann's dress is blue, it will be true in *w* if Ann's dress is red, etc. No difference in truth value at any index seems to arise. The two readings of the sentence, simply, seem to have the same truth conditions.[7]

In order to understand the crux of the trivialization problem, we need to acknowledge that there are two different uses of general terms in natural language. First, there are sentences whose subject matter are the things that exemplify the universal denoted by a general term (at the world of evaluation). The evaluation at a world of those sentences depends on the exemplars of the universal at the world in question. For instance, the sentence 'Tommy is a tiger' is true at a world *w* when Tommy is one of the things exemplifying the kind denoted by the term 'tiger' in *w*. We will say that general terms have an *exemplifier semantics* when they contribute to the truth conditions of the sentence to which they belong in this way (i.e. through their exemplars). When the general term occurs in predicate position in a sentence and it is attributed to an object, it always has an exemplifier semantics. But a general term can also appear in other syntactic positions and still have an exemplifier semantics, as long as the truth value at an index of evaluation depends on the things that instantiate the universal designated. As an example, consider the sentence 'tigers are cute': the truth of the sentence depends on whether exemplars of the species 'tiger' are or are not cute at the world of evaluation.

There is a second type of sentence, however, where the subject matter of the claim expressed by the sentence is the universal itself. The truth value of the sentence (relative to a possible world) is then determined only by the universal designated by the general term at the world of evaluation. For instance, when we say 'gold has atomic number 79' we talk about the substance gold, not just about its samples. We will say that an occurrence of a general term in a sentence has a *kind semantics* when the evaluation of the sentence depends on which universal is designated by the general term (at the index of evaluation).

As we saw before, LaPorte's examples show that there is no trivialization problem when general terms have what we have called here a *kind semantics*. The problem that LaPorte's argument still leaves unsolved can be formulated now as follows: there seem to be no examples of sentences in which a general term has an *exemplifier semantics* and whose truth value at some index of evaluation differs depending on whether the general term is interpreted rigidly or non-rigidly. [8]

On the face of it, this problem seems unsolvable. When a general term has an exemplifier semantics, the set of exemplars at a given index of the universal designated by the rigid reading of the term will coincide with the set of instances of the universal designated, at that same index, by the non-rigid reading. Recall sentence (1): the objects that have the property of being the colour of the sky at a given index—i.e. the objects that instantiate at that index the universal denoted by the rigid reading of 'the colour of the sky'—are the same objects that instantiate the colour designated non-rigidly by 'the colour of the sky' at that very index. Any object that exemplifies one universal will exemplify the other. Given that the truth value of a sentence in which a general term occurs in exemplifier semantics is ultimately determined by the things instantiating the universal designated, it seems natural to expect that the truth value at any index will be the same for rigid and non-rigid readings.

But the problem does have a solution. In fact, the tools for the solution were already available in Linsky 1984.[9] Consider the sentence:

> (2) It might have been the case that Loch Ness was not the actual colour of the sky.

Sentence (2) is true if, and only if, there is a possible world w in which Loch Ness is not the actual colour of the sky, i.e. when Loch Ness does not belong to the set of exemplars of the universal designated by 'the actual colour of the sky' at w. Let us consider first the non-rigid reading of 'the colour of the sky'. According to this reading, 'the colour of the sky' designates at each world the colour of the sky at that world; hence 'the actual colour of the sky' will designate rigidly the colour blue. On this reading sentence (2) is true just in case it might have been the case that Loch Ness was not blue. That is intuitively true, since there could be a world in which the sky was red, and Loch Ness would be red in that world, not blue.

Let us now move on to the rigid reading of 'the colour of the sky'. 'The actual colour of the sky' will designate rigidly the universal designated by 'the colour of the sky' at the actual world. But, since 'the colour of the sky', in its rigid reading, designates rigidly the universal that things instantiate when they have the property of being the same colour as the sky, 'the actual colour of the sky' will designate rigidly the same universal.[10] Sentence (2), under this interpretation, is true if, and only if, there is a world w in which Loch Ness is not one of the things exemplifying the property of being the same in colour as the sky in w. But this is intuitively false, since lakes reflect the colour of the sky, so they cannot fail to be the same colour as the sky. So sentence (2), under the rigid interpretation of 'the colour of the sky', is intuitively false.

We have presented an example of a sentence containing a general term that has an exemplifier semantics. The truth value of the sentence varies depending on whether the general term has a rigid or a non-rigid reading.

As we have pointed out, there is no difference in truth value nor in truth conditions between the rigid and the non-rigid reading of the general term in 'Loch Ness is the colour of the sky' and similar sentences. Nonetheless, there is a *semantic* difference between the two readings. The difference lies in the fact that the universals involved in the assignment of truth value, namely the *designata* of the two readings of the general term, are different.[11] Intensional operators, like the ones in sentence (2), allow us to bring that difference to the surface as a patent difference in truth value. There would be no difference in truth value between the two readings in (2) if there was no semantic difference between the two readings of 'Loch Ness is the colour of the sky'.[12] We conclude that there is no trivialization problem.[13]

3. THE OVERGENERALIZATION OF RIGIDITY

Some philosophers have objected that cashing out rigidity for general terms as sameness of designated universal across possible worlds results in an unwanted overgeneralization of the notion of rigidity, for only natural kind terms should be classified as rigid. Their claim is that, according to The View, 'bachelor' designates the universal *bachelorhood*, instantiated by bachelors, just like 'gold' designates the substance *gold* instantiated by samples gold. And if according to The View the latter is characterized as rigid, so will be the former:

> One thing to be said against the interpretation is that it would appear to have the consequence that a great many predicates will come out rigid . . . This is problematic, since Kripke wanted to distinguish natural kind predicates like *is gold* and *is a tiger* from ordinary descriptive predicates. (Soames 2002: 260)

Dan López de Sa presents the concern as follows:

> [The View's characterization of rigidity] would over-generalize, by covering predicates for artifactual, social, or evaluative properties, such as 'is a knife', 'is a bachelor', or 'is funny'. . . . And this despite the fact that the properties signified are, we may suppose, 'unnatural' (enough). According to the critics, this is an inappropriate over-generalization, as rigidity for predicates should apply only to predicates signifying natural (enough) properties—hence my labelling it 'the over-generalization problem'. (2008: 267)

Sensitive to this concern, some philosophers have endorsed an essentialist characterization of rigidity, according to which a general term is rigid just in case it expresses a property that is essential to its bearers, where an essential property is taken to be a property an object cannot fail to possess if it exists.[14] Using possible worlds semantics terminology, Michael Devitt characterizes rigidity for general terms as follows:

> a general term 'F' is a rigid applier iff it is such that if it applies to an object in any possible world, then it applies to that object in every possible world in which the object exists. (2005: 146)

Since natural kind terms are taken to express essential properties, on that account general terms such as 'tiger' and 'gold' are classified as rigid, and obviously 'philosopher' and 'bachelor' are not.

Now, the main issue here, from our point of view, is that overgeneralization should not be considered a problem. The View does classify terms such as 'pencil', 'bachelor', and 'philosopher' as rigid and, as we will argue, that is the correct way to classify those terms. But before we argue for that conclusion, there are four issues about the essentialist characterization that are worth discussing.

First, as many people have observed, the essentialist characterization of rigidity does not (at least in principle) seem to accommodate all of Kripke's remarks about rigid general terms. For one thing, 'yellow', for instance, does not express an essential property of the objects under its extension, and yet Kripke included colour adjectives among the list of rigid terms.[15] Another common complaint against the essentialist characterisation targets its excessive metaphysical commitments. As Devitt acknowledges: 'Clearly, if "*F*" is a rigid applier then any individual *F* must be essentially *F*. So the view that there are any such "*F*"s entails a fairly robust metaphysical thesis' (2005: 146). The concern here is not so much about essentialism *per se* as a metaphysical thesis; it is rather a worry as to whether what is primarily meant to be a semantic notion, a distinction that applies to language, should be characterized in such blatantly metaphysical terms.[16]

Second, Soames has argued that the essentialist strategy cannot account for the necessity of theoretical identifications, one of the allegedly crucial desiderata of a correct characterization of rigidity for general terms.[17] Our concerns about the essentialist strategy have nothing to do with this issue. In fact, although Devitt (2005) thinks that this is indeed a problem that the essentialist characterization—which he endorses—has to live with, Gómez Torrente (2006) has shown that the essentialist strategy can account for that desideratum.[18]

Third, in our view there are concerns to be raised about the satisfaction of the very first desideratum (see §1), for it is not entirely clear that the essentialist characterization of rigidity constitutes a natural extension of the notion of rigidity as it was defined for singular terms. On the face of it, the essentialist definition of rigidity for general terms seems to be a mirror image of the definition of rigidity for singular terms. Whereas in the case of singular terms we have

> a singular term *t* that designates object *o* is rigid just in case for all worlds *w*, such that *o* exists in *w*, *t* designates *o* in *w*,

in the case of general terms we have something like

> a general term *G* that applies to object *o* is rigid just in case for all worlds *w*, such that *o* exists in *w*, *G* applies to *o* in *w*.

However, it is not clear that the essentialist characterization satisfies the intuitive test of rigidity proposed by Kripke. When he introduces the notion of rigidity for singular terms, Kripke remarks:

> One of the intuitive theses I will maintain in these talks is that *names* are rigid designators. Certainly they seem to satisfy the intuitive test mentioned above: although someone other than the U.S. President in 1970 might have been the U.S. President in 1970 . . . no one other than Nixon might have been Nixon. . . . I will argue, intuitively, that proper names are rigid designators, for although the man (Nixon) might not have been the President, it is not the case that he might not have been Nixon. (1980: 48–9)

So, since

(3) Nixon could not have failed to be Nixon

is true no matter how we read the scope of the modal operator,

(4) Nixon could not have failed to be the president of the U.S.

is false, and

(5) The president of the U.S. could not have failed to be the president of the U.S

has one intuitively false reading, we conclude that 'Nixon' is rigid and 'the President of the U.S.' is not.

Generalizing the test, we could say that if a sentence of the form t_1 *could not have failed to be* t_2 (where t_1 and t_2 can be occurrences of the same or different terms) has at least one intuitively false reading, at least one of the terms t_1 and t_2 is not rigid. Thus, 'Nixon' and the supposedly codesignative 'the offspring of gametes ABC' are rigid, and 'the President of the U.S. in 1970' is not rigid.

When we apply the test to general terms, such as 'water', 'H_2O', and 'stuff that fills our lakes', we obtain:

(6) Water could not have failed to be water.
(7) Water could not have failed to be H_2O.
(8) Water could not have failed to be the stuff that fills our lakes.

Since (6) and (7) are intuitively true, 'water' and 'H_2O' should be classified as rigid. So far so good. But we think that when we reflect on the reasons that prompt us to say that (6) and (7) are true and (8) is false, it is easy to see that the intuitions at play have to do with what we take to be the nature of the *kind*, the substance Water, not with the properties of individual samples of water. The intuition that the substance could not have failed to be the very substance it is, and have the very molecular structure it has, is what is driving our intuitions about the truth value of (6) and (7).

In fact, one may be quite ready to declare (6) and (7) true, without having any clear intuitions about the essential properties of *individual* samples of stuff that happen to be water. The nature of a kind may not be an essence of the members or samples of the kind. Essences of kinds are properties or conditions that individual members or samples of a kind possess, or satisfy, by virtue of being members or samples of that kind. But those properties or conditions may fail to be essential to the individual members or samples of the kind.

The stuff in my glass has molecular structure H_2O in virtue of being water; and in every possible world that particular sample will have molecular structure H_2O *if* it is water. That much follows from (5). But it does not follow that the particular sample of water in question is itself *necessarily* water.

Similarly, when we consider sentences such as

(9) Tigers could not have failed to be tigers

and

(10) Tigers could not have failed to have DNA_{Tiger}

our tendency to regard those sentences as true (supposing that having a certain DNA is essential to being a member of the species Tiger) is driven by considerations about the species and what it takes to be a member of the species, not by strong metaphysical intuitions about the essential properties of members of the species, such as the individual tiger Tommy.[19]

The intuitions about essence that we appeal to when we consider sentences such as (6)–(10) surface because we are considering (6)–(10) as statements *about* kinds, not about members or samples of kinds. We are not judging those intuitively rigid general terms to be rigid appliers, nor are we concluding that they express properties essential to the objects or samples under their extension.[20]

When Soames discusses the essentialist approach to the rigidity of general terms, he is ready to accept the essentialist characterization as a natural extension of the notion defined for singular terms. His main objection to the approach, as we mentioned before, is based on the alleged difficulties of accounting for the necessity of theoretical identifications. But we think that Soames's—and the essentialist's—insistence in considering general terms only as predicates obscures the fact that the essential properties relevant in the assessment of sentences such as (6)–(10) are properties of the kind, or universal, not of its members or samples, and it does so by not recognizing that the general terms in those sentences are used to talk about a universal, not to attribute it. The distinction between those two uses of general terms, and the associated distinction between kind semantics and exemplifier semantics, is crucial here to understand what is driving the evaluation of sentences (6)–(10) and the judgments about the rigidity of the general terms they contain. The conclusion that terms such as 'tiger' and 'water' are rigid appliers is not grounded on the application of the general term version of Kripke's intuitive test for rigidity, and thus the essentialist cannot claim that his characterization of the rigidity of general terms is a natural extension of the notion of singular term rigidity, as required by the first desideratum.

Fourth, there is a metaphysical concern about the essentialist approach as it is stated. This view relies on the assumption that whether or not a property is essential is independent of which things bear the property, i.e. that the question about essence can be answered only by reference to properties ('is *being red* an essential property?'; 'is *being a tiger* an essential property?'). Now, we think one should not dismiss *by fiat* the possibility that some properties may well be essential to some bearers and non-essential to others. The view that some properties apply contingently to some objects and necessarily to others may be the wrong metaphysical picture, but it is not *obviously and trivially* the wrong metaphysical picture.[21]

For instance, *being red* is, we can safely assume, not essential to the walls of my house. But there may be objects or compounds in nature whose underlying, and essential, properties are responsible for emitting light or reflecting light in a certain way. And on some theories of colour, at least, this would mean that their being red is a property those objects simply cannot fail to have. If that were the case, it is hard to tell how the proponents of essentialist rigidity would want to apply the account. One alternative would be to define rigidity relative to the possessors of properties, so 'red' may be non-rigid as applied to the walls of my house and it may be rigid as applied to certain compounds in nature. Although this is not, by any means, a knockdown objection, we think that the need to make such a move pushes the essentialist characterization of rigidity further away from the realm of semantics.

These four issues are very important but, as we indicated before, our main concern about the essentialist characterization of rigidity lies precisely in what its proponents see as one of its advantages: the fact that terms such as 'bachelor', 'pencil', or 'philosopher' are, on the essentialist characterization, non-rigid. Quite the contrary, if we focus exclusively on *semantic* behaviour, all simple name-like general terms—'gold', 'water', 'bachelor', 'philosopher'—are semantically alike and should fall in the same category.

One of the main goals of *Naming and Necessity* is to highlight the radical difference between names and definite descriptions. The difference between the semantic behaviour of names and definite descriptions is a special case of a more general distinction between simple name-like expressions that apply purely by linguistic convention and expressions whose content determines their domain of application by virtue of the obtaining of certain facts or the satisfaction of certain (necessary or contingent) worldly conditions.[22] That is a difference that applies to singular and to general terms equally. On the Putnamian idealized model of introduction of a general term, we can say that someone *decided* to apply the word 'gold' to the original paradigms. But it is certainly not by decision that 'substance with atomic number 79' applies to them. Similarly, it is by decision, or by convention, that we use 'bachelor' to apply to unmarried adult males, 'physician' to people with a certain profession, and 'pencil' to a certain kind of tool. If we focus on *semantic behaviour*, 'tiger' and 'physician' should be in one category, 'substance with atomic number 79', and 'person who treats illnesses' in another. Simple, name-like general terms display a similar semantic behaviour and so they belong in the same semantic category, different from the category of complex general terms, precisely for the same reasons that proper names and singular terms with descriptive content belong in different categories.[23]

Thus, we think that the second desideratum that, allegedly, an adequate characterization of rigidity for general terms should satisfy is incorrect. General term rigidity should not be circumscribed to natural kind terms.[24] The fact that all simple general terms are characterized as rigid is, we think, not an undesirable overgeneralization of rigidity. It's just the way things should be.

4. RIGIDITY AND THE NECESSITY OF THEORETICAL IDENTIFICATIONS

The distinction between kind semantics and exemplifier semantics allows us also to address the discussion of the generally accepted desideratum that a satisfactory characterization of the notion of rigidity for general terms

> must play a role in explaining the necessity of true 'theoretical identification sentences' of the form [$\forall x(Gx \rightarrow Hx)$ and $\forall x(Gx \leftrightarrow Hx)$] containing rigid predicates that is analogous to the role played by the notion of a rigid singular term in explaining the necessity of true identity sentences of the form [$t = t'$]. (Soames 2002: 263)

Now, natural language sentences of the form *G is H*, *Gs are Hs,* or *a G is an H* can be interpreted as sentences in which the general terms in question have a predicative role and assert that objects (or samples) fall under the domain of application of *G* just in case they fall under the domain of application of *H*. The form of these sentences can be captured as $\forall x(Gx \leftrightarrow Hx)$.[25] In those cases, the rigidity of *G* and *H* does not entail the necessity of the true $\forall x(Gx \leftrightarrow Hx)$, for *G* and *H* may well rigidly designate two different universals that happen to apply to the same things.

On the other hand, those sentences can also be interpreted as expressing claims about the universals designated by the terms, claims that are true if the universal designated by *G* and *H* is the same, i.e. they can be interpreted in terms of kind semantics, a reading that can be captured as $G = H$. If G and H are rigid, those sentences are, if true, necessary, for G and H designate the same universal (represented as an intension) in every possible world.

Distinguishing kind semantics from exemplifier semantics allows us to see clearly that sentences of the form *G is H* , *Gs are Hs,* or *a G is an H*, if interpreted as claims whose truth depends on exemplars, *should not* be expected to be necessary on the basis of the rigidity of the general terms involved. Two general terms may rigidly designate different universals that happen to be actually instantiated by the same objects or samples.[26] But it is a mistake to think that sentences of the form *G is H*, *Gs are Hs,* or *a G is an H* can be interpreted only in terms of exemplifiers.

In fact, we think that theoretical identifications are *not*, and should not be conceived as, sentences of the form $\forall x(Gx \leftrightarrow Hx)$. 'Water is H_2O' or 'gold is the substance with atomic number 79' are not just universally quantified bi-conditionals; what makes them true is something over and above the fact that all the objects or samples that fall under the extension of one of the terms fall also under the extension of the other. Theoretical identifications are claims about kinds, about natural phenomena, species, and substances, and so their natural interpretation should be given in terms of kind semantics.[27]

When *G is H* is interpreted in terms of exemplifier semantics, it can be true and contingent—even when G and H are rigid general terms. In that case the interpretation of *G is H* in terms of kind semantics will be a falsehood. This is so because if $\forall x(Gx \leftrightarrow Hx)$ is contingent, G and H designate different intensions, different universals—in which case $G = H$ is false.

But if G and H are rigid and *G is H*, interpreted in terms of kind semantics, is true, it will also be necessary. The truth of the identification, so interpreted, entails its necessity, as per the desideratum.

There seem to be counterexamples to this claim; i.e. there seem to be sentences of the form $G = H$ that are true but contingent, even though G and H are rigid. If we are right, this should never be the case: identity sentences between rigid terms are either false or necessary. But a sentence such as 'renates are cordates', interpreted as an identity claim, appears to provide a counterexample. If the terms are interpreted as having kind semantics (as in 'a renate is a cordate'), it appears that both terms, arguably rigid terms, designate the same kind of thing, but they do so contingently.

We think that a moment's thought reveals that this is not a counterexample. To see why, we must pay heed to the connection between the kind and the exemplifier interpretations of an identification sentence. As Soames has noted (2002: 260), the truth of *G is H*, *Gs are Hs,* or *a G is an H* interpreted as expressing a claim about universals ($G = H$) will entail the necessity of *G is H*, *Gs are Hs,* or *a G is an H* interpreted as a claim whose truth depends on exemplars, a claim of the form $\forall x(Gx \leftrightarrow Hx)$. Observe that the sentence $\forall x(Cordate\ x \leftrightarrow Renate\ x)$ is intuitively contingent.

If that sentence is indeed contingent, there are possible worlds in which objects that fall under the universal designated by 'cordate' do not fall under the universal designated by 'renate'.[28] But that just means that those are two different universals, two different kinds of things that happen, but just happen, to be exemplified by the same objects. So *Renate* = *Cordate* is not a counterexample, because it is false. There are no true but contingent genuine theoretical identifications involving rigid terms. To sum up, *G is H*, when interpreted in kind semantics, satisfies the desideratum: if *G* and *H* are rigid, *G is H* is necessary, if true. If *G is H* is interpreted in terms of exemplifier semantics, it does not satisfy the desideratum. And it should not do so.

It is worth reflecting further on the connection between the kind semantics interpretation and the exemplifier semantics interpretation of the rigid *G is H*. It is not just that, as Soames has noted, the truth of $G = H$ will entail the necessity of $\forall x(Gx \leftrightarrow Hx)$ (as well as the necessity of $G = H$ itself). It is also the case that, if $\forall x(Gx \leftrightarrow Hx)$ is necessary, $G = H$ will also be true (and consequently necessary). For, if $\forall x(Gx \leftrightarrow Hx)$ is necessary, the exemplars of *G* coincide with the exemplars of *H* in every possible world, and therefore the intension that represents the kind designated by *G* is the same as the intension that represents the kind designated by *H*. Now, this may initially seem to be an unwelcome result. But we think that this is as it should be.

Suppose that a group of explorers in some distant planet discover animals whose young nurse the milk of their mothers and classify them as 'noors'. In that planet they observe also animals that have hair on their bodies, and classify them as 'hayrs'. Suppose now that the sentence 'noors are hayrs' (interpreted as a sentence of the form $\forall x(Gx \leftrightarrow Hx)$) is true. The two terms are rigid. As we have argued, we should not, just on the basis of the rigidity of the general terms, expect 'noors are hayrs' to express a necessary truth. Were there to exist an essential connection between having hair and nursing, the sentence would express a necessary truth. Were there such an essential connection, 'hayrs' and 'noors' would be general terms designating the same kind of individual. And on the kind semantics interpretation, 'noors' and 'hayrs' would then designate the same kind, represented by one and the same intension. Two general terms *G* and *H* may single out the same substance or species even though they may be associated with different criteria of classification. If it turns out that the objective similarities that are responsible for objects or samples falling under the terms are essentially connected, then the corresponding statement of the form $\forall x(Gx \leftrightarrow Hx)$ is necessary, for *G* and *H* identify the same kind. And since the two terms designate the same kind of thing, $G = H$ will also be necessarily true.

The reason why this seems to be an unwelcome result rests, we believe, on a confusion, although the confusion rightly points to what is, in fact, a general shortcoming of possible worlds semantics that affects any explanation of any semantic phenomenon given by appeal to intensions.

Prima facie, it seems that there may be sentences of the form $\forall x(Gx \leftrightarrow Hx)$, with *G* and *H* rigid, that are necessarily true even though the terms *G* and *H* do not intuitively designate the same kind.[29] Terms such as 'shaped' and 'extended' might, perhaps, be an example of this.

But we think that the appearance is an illusion that rests on an underlying confusion between, on the one hand, the universals or kinds *designated* by general terms and, on the other hand, the criteria of identification of those kinds, or the properties expressed by the kind terms in question.

If exemplars of *G* and exemplars of *H* are the same individuals or samples as a matter of necessity, *G* and *H* identify the same kind of thing, or the same kind of stuff. Now, it may well be that the criteria by which things or samples of that kind are identified are different; it may well be that the *properties* that those things or samples satisfy, properties in virtue of which the objects or samples are members of the kind, are different properties. In that case, *G* and *H* are associated with or, if one wishes to put it this way, *express* different properties. 'Noor' and 'hayr' and perhaps 'shaped' and 'extended' express different properties; at the very least, they are associated with intuitively different conditions of membership to a kind. But none of this entails that they identify different kinds.

Now, something that may be in part responsible for the mistaken assumption that each term in the pairs 'noor' / 'hayr', 'shaped' / 'extended' *designate* different kinds may well be the judgment that each of them *expresses* a different property. And the intuition that those properties are indeed different may be responsible for the mistaken assessment of the identity sentences as false. Now, it is true that in possible worlds semantics the intuitive difference between the property of being a noor and the property of being a hayr (and the intuitive difference between the property of being shaped and the property of being extended) cannot be represented. [30] Intensional semantics cannot distinguish between the properties expressed by terms that designate the same universal, simply because the *intensions expressed* by those terms (functions from indices to the *intensions* the terms *designate*) are represented by the same functions. Possible worlds semantics has, indeed, its limitations; but the shortcomings are the same across the board: any two expressions (singular terms, general terms, sentences) that co-designate in every possible world, because of logical or metaphysical equivalence, express the same intension, so this is not a problem that afflicts specifically and especially the conception of kind terms as designators.

A more fine-grained semantical apparatus may be deemed preferable, for instance, to represent propositions or beliefs. But in the specific case that concerns us here, namely universals designated by general terms, the level of representation provided by intensional semantics gives us all we need. A significant part of scientific practice consists in finding accurate ways of characterizing substances and kinds via the discovery of necessary ties between properties. Even if we had a more fine-grained representation of the properties expressed by general terms, we would still need a way of representing classes of properties that are nomologically connected. Intensional semantics allows us to do that by making the general terms that express those different properties designate the same intensional function. The representation of kinds as intensional functions is valuable independently of the expressive limitations of intensional semantics.

5. CONCLUSIONS

Our purpose here has been to defend The View from its most damaging objections. As we have shown, The View does not trivialize the notion of rigidity for general terms. There are sentences whose truth value and truth conditions differ depending on whether a general term occurring in them is interpreted as rigid or as non-rigid, and that difference arises in sentences where the general terms are used non-predicatively, to make claims about the universals they designate, as well as in sentences where the general terms are used predicatively.

The View does, however, overgeneralize; i.e. it classifies as rigid general terms such as 'philosopher' or 'pencil'. But, as we have argued, that is precisely what an adequate characterization of the notion of rigidity for general terms *should* do. Finally, The View gives an account of what theoretical identifications are. Theoretical identifications are not just universally quantified bi-conditionals; they are *identity* sentences involving general terms, sentences that identify a kind designated in two different ways. And it is precisely the distinction between universalized bi-conditionals, sentences where the general terms are used predicatively and with exemplifier semantics, and identity sentences where the general terms are interpreted in terms of kind semantics, that allows us to explain why true genuine theoretical identifications are necessary.[31]

NOTES

1. The trivialization problem is raised by Schwartz (2002) and Soames (2002: 261).
2. See Soames (2002: 263) for a statement of the desiderata. The alleged failure of proposed characterizations of the rigidity of general terms to satisfy one or more of the desiderata leads Soames and Schwartz to abandon the project of defining general term rigidity. For similar reasons, Glüer and Pagin (forthcoming) also propose to abandon the project of extending the notion of rigidity to general terms in favour of looking for semantic properties held in common by natural kind terms.
3. This problem is raised by Soames (2002: 259–60). The term 'overgeneralization' is López de Sa's (2008).
4. See Soames 2004.
5. We think it is very important to distinguish kinds from properties, so we distinguish the substance gold from the property of being gold, the colour blue from the property of being blue. More about this in §4.

6. It could be argued that the rigid reading is very unnatural; terms such as 'the colour of the sky' and 'beverage requested by Uncle LaPorte at Super Bowl parties' do not designate *kinds* of things, quite simply because there are no such properties in the world as being the colour of the sky or being the beverage requested by this or that person at certain parties. But, from a purely semantic point of view, there is in principle no reason to deny legitimacy to such properties, to the kinds of things that instantiate them, and to the rigid readings of terms that designate those kinds. As Schwartz (2002: 268) has pointed out: ' . . . kinds may have an important role in our common sense understanding of the world and even in science, but they don't have a metaphysical status that is useful to formal semantics. . . . [P]roperties are not limited to robust things like causal powers—they are simply sets of actual and possible individuals and for every such set there is a property'. Besides, it is arguable that there are some properties that correspond to the rigid readings of some *prima facie* non-rigid general terms. In the particular case under consideration here, the property of being the colour of the sky is, arguably, a natural property that some surfaces have by virtue of their capacity to reflect colours. If so, 'the colour of the sky' can be interpreted rigidly as designating a natural kind, the kind of thing that reflects the colour of the sky.
7. See Martí and Martínez-Fernández 2007 for a step-by-step explanation, within the framework of possible worlds semantics, of how the problem arises.
8. In his response to a proposal by Martí (2004), López de Sa (2007) argues that the problem of trivialization persists when one considers sentences such as (1), for reasons similar to the ones pointed out here. Curiously, in his own proposed characterization of the rigidity of general terms, López de Sa gives no example where any difference in the truth conditions of sentences containing general terms can be detected depending on whether the terms are interpreted rigidly or non-rigidly.
9. When Linsky's paper was published the trivialization objection had not been posed as an objection in these terms.
10. The case here is parallel to the singular term case: 'the actual successor of eight' rigidly denotes the same number as 'the successor of eight'. If 'the offspring of gametes *a* and *b*' designates Cicero rigidly, so does 'the actual offspring of gametes *a* and *b*'. 'Actual' is idle when applied to a rigid term, be it singular or general.
11. Some 20 years after the 1984 paper, Linsky points out that the difference can only show up under the scope of certain intensional operators, precisely because although the truth condition of a sentence like (1) is the same under the two readings "the way this condition is determined is different in the two cases" (Linsky 2006: 659–60).
12. If there is no semantic difference at all between S and S', adding modal operators is not going to create a difference.
13. In our 2009 manuscript we present a formal analysis of our solution to the trivialization problem using the tools from Linsky 1984.
14. See Devitt 2005 and Gómez Torrente 2006.
15. Mario Gómez Torrente (2006) points out that, while adjectival uses of colour predicates are not rigid on the essentialist characterization, substantival uses are. In his view, the essentiality of the nouns underlies the intuitions concerning the necessity of certain theoretical identifications. Gómez Torrente reports that he is developing an account of colour language of which the claims just stated are consequences. See the following for our discussion of theoretical identifications.

16. For discussions of these two worries, see Inan 2008 and Martí 2004. Arguments against grounding linguistic distinctions on metaphysical theses have their origins in Salmon 1982.

17. See Soames 2002: 254–9.

18. Gómez Torrente argues that if rigidity is understood as obstinate rigidity and quantification is possibilist (two, arguably, plausible assumptions when one considers *theoretical* identifications), the necessity of true identifications involving general terms follows unproblematically.

19. Ilhan Inan is also concerned about the failure of essentialist characterization to satisfy the first desideratum: 'In ordinary discourse we could . . . [say] 'the colour blue might not have been the colour blue' . . . and 'blue things might not have been blue' . . . It seems to me clear that when Kripke introduced this test to decide whether a particular general term is rigid, he meant the former and not the latter . . . [It] would be incorrect to say 'Truth might not have been truth', but correct to say 'true propositions might not have been true' for at least some propositions. From the latter claim, nothing follows about the rigidity of the term; it is the former that is important' (2008: 219). Inan argues also that this confusion underlies the claim that artifact terms such as 'pencil' should not be rigid (see Inan 2008: 219–20). For a discussion of artifact terms, see also Schwartz 2002, Devitt 2005, and Glüer and Pagin (forthcoming). At the end of this section we offer a general argument, on purely semantic grounds, against the claim that only natural kind terms should be taken to be rigid.

20. This is not to deny that individual members of the species Tiger are essentially tigers, or that individual samples of water are essentially H_2O. The point here has to do with the grounds on the basis of which we judge the truth value of sentences such as (6)–(10).

21. Jessica Wilson defends the relativity of essence: ' . . . whether a property is contingent or essential is relative to what substantial particular has the property (being charged may be essential to an electron, but inessential to me)' (Wilson unpublished: 2).

22. See Martí 2004 for further discussion of this issue.

23. Certainly, rigidity is too coarse a semantic sieve: some definite descriptions get still classified together with proper names, and similarly, some complex general terms will be classified also as rigid. But the problem with the essentialist interpretation is that it gets the classification even more wrong. It lets in some terms with descriptive content and it also leaves some name-like general terms out.

24. See Wikforss (in this volume) for a entirely different argument—a descriptivist argument—for the claim that simple natural kind terms are not semantically different from other simple general terms.

25. We are, for the moment, leaving aside the discussion of sentences such as 'cats are mammals' and 'blue is a colour', i.e. sentences that are neither statements of identity nor interpretable as universalized biconditionals.

26. See Donnellan 1983 for a discussion casting doubts on the assumption that the extension of the notion of rigidity to general terms should give rise to what he calls 'exotic necessary truths'.

27. Notice that sentences such as 'water is H_2O' or 'a bachelor is an unmarried adult male' can be interpreted in two ways, depending on whether the terms have kind or exemplifier semantics. The source of the ambiguity is, as one would expect, the copula, which can express identity or coextensionality. Other interesting issues arise when one considers sentences such as 'gold is a metal', which can also be interpreted as claims about the universal (on kind semantics) or as claims about the samples of the substance (as a universally quantified conditional). The discussion of these cases is beyond the scope of this paper.

28. We are assuming here that it is possible to have a heart that pumps blood without having a filtering device such as kidneys.
29. This is the mirror image of the 'renates' and 'cordates' case.
30. We are distinguishing here between properties and other universals such as kinds because we think that a kind can be identified via two distinct essential properties. Linsky (1984: 262) treats kinds as special properties, the property of being a member of the kind. We think that our stance is more plausible because it acknowledges, even if it cannot represent it, that a kind may have two different properties that are nomologically connected.
31. We thank the audiences of the Nature and Its Classification conference held in Birmingham in November 2007 and the 37th Annual Meeting of the Society for Exact Philosophy, held in Alberta, Canada, May 2009. We are also grateful to Ilhan Inan, Bernard Linsky, Jeff Pelletier, and Jessica Wilson for helpful comments and discussion. The research for this paper has been partly funded by grant 2008-FFI04263 awarded by the Spanish Ministerio de Ciencia e Innovación.

REFERENCES

Devitt, M. (2005) 'Rigid application', *Philosophical Studies*, 125: 139–65.

Donnellan, K. (1983) 'Kripke and Putnam on natural kind terms', in C. Ginet and S. Shoemaker (eds), *Knowledge and Mind*. Oxford: Oxford University Press.

Glüer, K. and Pagin, P. (forthcoming), 'General terms and relational modality', *Noûs*.

Gómez Torrente, M. (2006) 'Rigidity and essentiality', *Mind*, 115: 227–59.

Inan, I. (2008) 'Rigid general terms and essential predicates'. *Philosophical Studies*, 140: 213–28.

Kripke, S. (1980) *Naming and Necessity*. Cambridge, MA. Harvard University Press.

LaPorte, J. (2000) 'Rigidity and kind', *Philosophical Studies*, 97: 293–316.

Linsky, B. (1984) 'General terms as designators', *Pacific Philosophical Quarterly*, 65: 259–76.

———. (2006) 'General terms as rigid designators', *Philosophical Studies*, 128: 655–67.

López de Sa, D. (2007) 'Rigidity, general terms, and trivialization', *Proceedings of the Aristotelian Society*, 107: 117–23.

———. (2008) 'The over-generalization problem: predicates rigidly signifying the "unnatural" ', *Synthese*, 163: 263–72.

Martí, G. (2004) 'Rigidity and general terms', *Proceedings of the Aristotelian Society*, 104: 129–46.

Martí, G. and Martínez-Fernández, J. (2007) 'General terms and non-trivial rigid designation', in C. Martínez, J. L. Falguera, and J. M. Sagüillo (eds), *Current Topics in Logic and Analytic Philosophy*, Santiago de Compostela, Universidad de Santiago de Compostela, 103–16.

———. (2009) 'General terms, rigidity and the trivialization problem'. Manuscript.

Schwartz, S. (2002) 'Kinds, general terms and rigidity: a reply to LaPorte', *Philosophical Studies*, 109: 265–77.

Soames, S. (2002) *Beyond Rigidity*. New York. Oxford University Press.

———. (2004) 'Replies to Gómez Torrente and Ezcurdia', *Critica*, 36: 83–114.

Wilson, J. (unpublished) 'Trope determination and contingent characterization'. Manuscript.

4 Are Natural Kind Terms Special?

Åsa Wikforss

1. INTRODUCTION

These days we speak freely of 'natural kind terms', indicating that they constitute a special, semantic category of terms. It was not always so. Indeed, prior to Putnam and Kripke's writings from the 1970s, the label 'natural kind term' seems not to have been employed at all. In his well-known paper 'The Analytic and the Synthetic' (1962), Putnam does set out to draw some distinctions among the general terms that he takes to be of semantic significance. In particular he wishes to distinguish so-called one-criterion terms, such as 'bachelor', from cluster terms, such as 'man' or 'crow', where the meaning is given by a cluster of associated properties, none of which are immune from revision. Among the cluster terms, Putnam also suggests, there is a set of terms of special interest to science, the 'law-cluster terms', such as 'energy'. These are set apart by what goes into the cluster, in particular laws and general principles. In his paper 'Is Semantics Possible?' (1970), however, natural kind terms appear on the scene. As in the earlier paper, Putnam draws a distinction among the general terms between one-criterion terms and others, but he now drops the talk of law-cluster terms in favour of that of natural kind terms:

> A natural kind *term* . . . is a term that plays a special kind of role. If I describe something as a *lemon*, or as an *acid*, I indicate that it is likely to share certain characteristics (yellow peel, or sour taste in dilute water solution, as the case may be); but I also indicate that the presence of those characteristics, if they are present, is likely to be accounted for by some 'essential nature' which the thing shares with other members of the kind. What the essential nature is is not a matter of language analysis but of scientific construction. (1970: 140)

Putnam's formulation captures the central elements of the contemporary notion of a natural kind term: the idea that these terms pick out not superficial, macro-level properties, but underlying, essential properties the nature of which it is up to science to establish. The notion took on quickly, in particular after the appearance of Kripke's *Naming and Necessity* (1980) in 1972. By suggesting that natural kind terms, like names, are rigid designators, Kripke seemed to provide the semantic tools for separating out the natural kind terms from other kind terms. The suggested candidates included a rather diverse set of terms: mass terms ('water', 'gold'), count nouns ('tiger', 'whale') as well as adjectives ('hot', 'loud'). What sets these terms apart, Kripke suggested, is their 'name-like' semantic behaviour. In addition, Putnam presented his Twin Earth thought experiment which was also taken to indicate that natural kind terms are set apart: the meanings of these terms are not 'in the head', but have to be given an externalist account (1975a). Thus, the idea emerged that natural kind terms are semantically special.[1]

This idea has stayed with us. In the contemporary discussion the notion of a natural kind term plays a prominent role. At the same time, it is recognized that things are more complicated than initially thought. For instance, Kripke and Putnam's discussions were based on rather naive metaphysical (and scientific) assumptions about natural kinds, assumptions that have since been challenged. It also turns out that the semantic issues are less straightforward than assumed—in particular, it is far from clear what it might mean to say that a *kind* term is rigid. Strikingly, however, these worries have not done much to undermine the assumption that natural kind terms form a special semantic category. Indeed, a number of people have recently written on natural kind terms with the mission of spelling out exactly why they are, after all, special.[2] Although the resulting suggestions vary a great deal, it is agreed that there is *something* special about these terms.

In this paper I try to shake that confidence. I argue that the time has come to question the assumption that natural kind terms form a separate semantic category among the kind terms.[3] The semantic and metaphysical difficulties noted in the contemporary debate should be taken seriously, and cannot be dismissed as mere wrinkles. Indeed, a serious problem with the contemporary discussion is the assumption that the semantics of natural kind terms can be conducted independently of the metaphysical and scientific issues surrounding these kinds. It may of course be suggested that this is as it should be: We should be able to do our semantics independently of metaphysics and science. However, in the case of natural kind terms, at least, this is a dubious policy. The reason is simple: if it is distinctive of these terms that they pick out natural kinds, the essences of which have to be decided by science, then the metaphysics and science of natural kinds cannot be ignored.

When discussing the special status of natural kind terms, it is important to separate two ways in which a term may be semantically special: the term may have a special type of *semantics*, or the term may have a special *metasemantics* or foundational semantics—that is, there may be something special about how the meaning or semantic content of these terms are *determined*. For instance, when it is claimed that natural kind terms are special because they are rigid, this is a claim of the first sort, whereas when it is claimed that natural kind terms are special because their meaning is determined externalistically, it is a claim of the second sort. Although often conflated in the debate, the claims are quite independent of one another—for example, it may be that the meaning of natural kind terms is determined externalistically, and that this sets natural kind terms apart from other kind terms, even though natural kind terms have the same type of semantic content as other kind terms.[4] I shall therefore discuss the semantic and the metasemantic issues separately. In section 3, I discuss the two leading attempts to single out natural kind terms from other kind terms at the level of semantics: by appealing to rigidity and by appealing to non-descriptionality. Section 4 discusses the proposal that natural kind terms have a special metasemantics.

First, however, a prior question has to be addressed: What determines whether a term is a natural kind term in the first place?

2. WHICH TERMS ARE NATURAL KIND TERMS?

In the case of names, it seems that we have a fairly clear pre-theoretical conception of how to identify them: a term is a name, roughly, if it is used to refer to a particular individual. Thus, the suggestion that names constitute a special semantic category can draw on the intuition that names are easily separated from other terms of the language. In the case of natural kind terms, matters are more complicated since natural kind terms form a subgroup among the larger group of kind terms. We therefore need to know how to separate out this subgroup.

According to one proposal, whether or not something is a natural kind term depends on our *semantic intentions*. Determining whether a term is a natural kind term is therefore something that can be done *a priori*, by consulting one's intentions. For instance, there is the idea that what is distinctive of natural kind terms is that they (like indexicals) are associated with a semantic rule (character) that serves to fix the property (and hence the content) picked out by the term in a given context.[5] The rule in question utilizes the macrophysical properties associated with the term, but the essential property picked out is assumed to be an underlying, non-manifest one.[6] This type of proposal is driven by the conviction that there must be something about our use of the term, something that is *a priori* available and makes it the special kind of term it is (just like there is something about our use of an indexical, that is, *a priori* available and sets it apart from other terms). Another apparent advantage is that the proposal avoids making semantics hostage to metaphysics: since the semantic status of a term as a natural kind term depends wholly on the speaker's intentions, we can do the semantics of these terms, it seems, without having to worry about the metaphysics (and science) of natural kinds.

However, while it can be known *a priori*, on this view, whether a term is a natural kind term, it cannot be known *a priori* whether it succeeds in picking out a natural kind. The question therefore arises how to account for cases where a purported natural kind term fails to pick out a natural kind. Such examples are legion, at least if we rely on the microstructural conception of natural kinds (as people in the Kripke-Putnam tradition tend to), according to which the essential properties of natural kinds are microstructural: in many cases of purported natural kind terms there is not a unifying microstructural property of the sort required. A well-known example is 'jade' but similar problems arise for 'sugar', 'air', 'sand', and a multitude of names for plants, animals, and diseases.[7] And it may of course turn out that we are mistaken about there being such a property even in the case of 'water'.[8] On the proposal under consideration it would seem that in such a scenario the rule fails to pick out a property and thus determine a content, and hence that statements involving the term in question would fail to express anything. This type of problem is not unique to natural kind terms of course; we are familiar with the problem of non-referring names. However, it would seem to constitute a more serious problem in the case of natural kind terms. After all, terms such as 'water' and 'air' play an absolutely central role in our lives and practices and we simply cannot accept a theory that has the implication that all such discourse lacks content if it turns out that the term in question fails to pick out a unified, underlying structure. As Scott Soames puts it, this conclusion 'seems harsh' since 'during the period in question, speakers used sentences containing the term to convey lots of information' (2002: 281–2).[9]

Of course, the *a priori* proposal is not wedded to any particular conception of natural kinds, such as the microstructural conception. However, it is clearly committed to *some* such conception. After all, whether a purported natural kind term fails to refer, on this view, depends both on how the relevant intention is spelled out and how the specific natural kinds (species, chemical kinds, minerals, plants, etc.) are to be understood. Here, it should be noted, there is an obvious danger. If the speaker's intention (the reference-fixing rule) is vaguely specified, then there will be massive reference failures. The intention, notice, is supposed to carry with it a uniqueness requirement: the term is intended to pick out *the* underlying kind, and since every object is an instance of an infinite number of kinds (including natural kinds), the uniqueness requirement will most likely fail if one sticks with vague conceptions such as 'natural kind' or 'underlying property'. One might try to remedy this by appealing to more specific notions, such as 'chemical composition', 'species', or 'mineral'. This causes obvious troubles if we look to the historical use of these terms, since these notions were not available until the development of modern science.[10] But the question also arises how the ordinary concepts of species, minerals, chemical kinds, etc., correspond to those of science. If they do not correspond, it would seem that, again, there is a danger of massive reference failure—simply because the ordinary conceptions of essential properties do not line up with anything recognized as such by science.[11]

The problems caused by possible reference failures have led many people to abandon the *a priori* proposal in favour of an account according to which the status of a term as a natural kind term is a wholly *a posteriori* matter. On this view, whether a term is a natural kind term is not determined by our semantic intentions, but by the external world itself.[12] Although we believe 'water' to be a natural kind term, this belief will be mistaken if in fact 'water' fails to pick out a natural kind. In such a scenario the proper conclusion is not that the term fails to refer but, simply, that the term is not a natural kind term.

Now, if a term is a natural kind term only if it in fact picks out a natural kind, it becomes rather obvious that the metaphysical issues cannot be avoided if we are interested in saying something of interest about natural kind terms. What is needed, quite clearly, is a theory about what separates natural kinds from other kinds in order to determine which terms are natural kind terms. Unfortunately, there are several competing theories (the microstructural theory, the causal homeostasis account, promiscuous realism, and so on) and, depending on which theory one adopts, one will draw the distinction between natural kind terms and other kind terms differently.[13] On some accounts the natural kinds will be few (including, for instance, chemical kinds but not biological kinds) and hence the class of potential natural kind terms rather limited, whereas on less restrictive accounts this class will be rather large (including not just 'water' and 'gold' but also 'tiger', 'tree', 'pain', 'blue', and, even, 'capitalism'). And, clearly, it has to be assumed that there *are* natural kinds, that natural kinds are distinct from other kinds, or else the class of natural kind terms will be empty.[14]

Moreover, on this view whether a term is a natural kind term becomes not only *a posteriori* but something that cannot be established prior to detailed scientific investigations. For instance, assuming that the microstructural conception of natural kinds is correct, determining whether 'water' is a natural kind term requires knowing something about the underlying structure of the liquid and that requires developed scientific theory and methods of empirical investigation.

In itself, this need not be problematic. It is perfectly innocuous to say that natural kind terms are terms that pick our natural kinds (just as artifact terms pick out artifacts and functional kind terms functional kinds) and, hence, that we typically do not know whether a term is a natural kind term prior to scientific investigations. However, it should be stressed what follows if we combine this idea with the claim that natural kind terms form a special semantic category: The upshot is that there is a category of terms that is special from a semantic point of view, even though identifying this category depends on the development of sophisticated empirical theories, such as contemporary chemistry or evolutionary theory. Consequently, which terms have this special semantic character (if any) cannot be known independently of detailed scientific investigations.[15]

There are therefore two competing conceptions of what makes a term a natural kind term in the first place: what I have called the *a priori* and the *a posteriori* proposal. Both illustrate the interaction of the semantic and metaphysical issues, although in different ways. The *a priori* proposal allows us to distinguish the natural kind terms prior to scientific investigations and in that sense ducks the metaphysical (and scientific) issues. However, the latter issues return once the question is raised whether these terms succeed in picking anything out. Since natural kind terms, on this view, pick out the kind intended, there is the danger of a mismatch between intention and world, allowing for massive failure of reference. The *a posteriori* proposal avoids this difficulty by deferring to nature, as it were: letting nature, rather than the speaker's intentions, determine whether a term is a natural kind term. However, if one takes natural kind terms to be semantically special, this also means that the *semantics* of these terms is deferred to nature. Prima facie, this is a rather startling suggestion. Just how startling will depend on which semantic feature one takes to be the distinguishing characteristic of natural kind terms—to which I now turn.

3. THE SEMANTIC CONTENT OF NATURAL KIND TERMS

3.1 Rigidity

The first attempts to separate out natural kind terms from nominal kind terms turned on the idea that the latter terms can, while the former cannot, be given analytic definitions spelling out necessary and sufficient conditions.[16] However, the proposal stands and falls with the assumption that other kind terms can be given analytic definitions. Even if it is granted that Putnam's 'one-criterion terms' lend themselves to such definitions, most kind terms are not like that. For instance, we would be hard pressed to find plausible necessary and sufficient conditions for terms such as 'sand', 'tree', and 'mud'. Indeed, it is precisely considerations of this sort that led people, including Putnam himself, to abandon traditional versions of descriptivist theories in favour of the cluster theory.

With the publication of *Naming and Necessity,* a more comprehensive attack on descriptivist theories was launched—an attack, moreover, that seemed to provide a clear sense in which natural kind terms are semantically special. Kripke not only provided a variety of arguments, most famously modal arguments, against descriptivist accounts of names; he also suggested that natural kind terms are closely related to proper names. Names, according to Kripke, are rigid designators: they designate the same object in every possible world (where the object exists). Descriptions, such as 'the president of the United States', are not rigid designators; hence, the content of a proper name cannot be understood in terms of such descriptions. Similarly, Kripke argued, natural kind terms are rigid designators and cannot be given a descriptivist account either, not even along the lines of the cluster theory (1980: 116–43).

Although the notion of rigidity is relatively clear in the case of proper names, however, it is not clear how it is to be understood in the case of kind terms. Kind terms typically function as predicates and so one would have to explain how the notion of rigidity applies to predicates. Moreover, even if kind terms are construed as a form of singular term, it is much disputed what it is that they designate. The question of how to extend the notion of rigidity to kind terms was raised in the 1980s and 1990s,[17] and it has come to be hotly debated the last few years as a result of Soames's influential book on Kripke (2002). Soames sets up three requirements on an interesting notion of kind term rigidity: it must be a natural extension of the notion of rigidity defined for singular terms; it must single out the natural kind terms; it must play a role in explaining the necessity of true theoretical identity sentences (2002: 263). Soames considers a number of attempts to define a notion of rigidity for kind terms that meet these criteria, but argues that they all fail. He concludes that the notion of rigidity does not apply to kind terms.

The most obvious way of meeting the first requirement is to construe natural kind terms as a form of singular term. This, also, is the strategy employed by a number of writers. It is granted that natural kind terms also function as predicates, serving to classify objects, but the hope is that the singular term usage and the predicate usage can be shown to be appropriately related.[18] According to this view, then, a kind term **t** is rigid iff there is a unique property which it stands for that determines its extension at each possible world (Soames 2002: 250).[19] This also promises to meet Soames's third requirement of explaining the necessity of theoretical identity sentences, such as 'Water is H_2O'. Assuming that such sentences are construed as proper identities, the fact that the terms flanking the identity sign are rigid designators ensures that the sentence is necessary, if true.

However, Soames rejects this proposal on the grounds that it fails to meet the second requirement, that of singling out the natural kind terms. 'Water' picks out the same unique property in every world but so do 'philosopher' and 'chair'. It follows, it would seem, that no distinction at all could be drawn between rigid kind terms and non-rigid ones. This problem has come to be called the 'triviality problem', and it has elicited several responses. Joseph LaPorte, for instance, has suggested that although terms such as 'philosopher' and 'chair' designate the same abstract kind in every possible world, there are kind designators that do not and are, in that sense, non-rigid: for instance, 'The insect species that is typically farmed for honey'.[20] Genoveva Martí, similarly, has argued that a distinction can still be drawn between simple, 'name-like' general terms (such as 'water', 'yellow', and 'philosopher') and complex general terms (such as

'Mary's favourite colour'). The former terms are rigid in that they designate the same property in every possible world. The latter terms, however, are normally used non-rigidly. Thus, 'yellow' designates the same property in every possible world, whereas 'Mary's favourite colour' designates different colours in different worlds. Complex general expressions *could* be used rigidly (as when 'Mary's favourite colour' designates the higher order property of being Mary's favourite colour), although this is less common, whereas simple general terms only have the rigid use.[21]

Now, it should be noted that there are, in fact, two triviality problems. First, there is the concern that all kind designators will come out as rigid, thus allowing no distinction between expressions such as 'water' and 'Granny's favourite drink'. Second, there is the concern that all simple general terms, natural and artificial kind terms alike, will be rigid.[22] Martí's and LaPorte's proposals address the first triviality problem but not the second one—that of separating out the natural kind terms among the rigid kind designators.

They are not unaware of this. LaPorte explicitly rejects the assumption that rigidity can be employed to show that there is an interesting distinction between natural kind terms and artificial kind terms:

> I cannot agree with various suggestions, then, that non-natural or nominal kind terms stand in contrast to natural kind terms over rigidity, just as descriptions like 'the inventor of bifocals' contrast as non-rigid designators with names, such as 'Ben Franklin'. The proper contrast over rigidity is that between non-rigid descriptions for kinds (either natural or artificial), on the one hand, and rigid names/descriptions for them on the other. (LaPorte 2000: 299)

Martí, similarly, denies that her notion of rigidity can be employed to separate out the natural kind terms. Indeed, Martí takes it to be an advantage of her position that it does *not* distinguish the natural kind terms from other simple general terms, such as 'philosopher': 'Simple, name-like general terms display a similar semantic behavior and so they belong in the same semantic category, different from the category of complex general terms, precisely for the same reasons that proper names and singular terms with descriptive content belong in different categories' (2004: 133–4).[23]

According to this line of argument, therefore, there *is* an interesting notion of rigidity that applies to kind terms only it does not serve to single out the natural kind terms from other, name-like kind terms. That is, Soames's second requirement is rejected.[24] Indeed, the upshot seems to be that natural kind terms and (simple) non-natural kind terms are more *closely* related than previously thought. In the case of Martí, at least, the notion of kind term rigidity goes hand in hand with non-descriptionality, and this suggests that not only natural kind terms but also other (simple) kind terms are non-descriptional. This conclusion is explicitly drawn by

Nathan Salmon. If simple kind terms ('water', 'blue', 'bachelor') all turn out to be rigid designators, he suggests, it follows that a term such as 'bachelor' functions like a logically proper name, rather than a description, of the gendered marital-status category *Unmarried Man*: 'If that is how it does function, then its rigidity is *de jure* and, contrary to the common view, it is not strictly synonymous with the corresponding description, even though it is closely tied to the description—as the name "Hesperus" is closely tied to some description of the form "the first heavenly body visible at dusk . . ." ' (2003: 486–7).

On this view, therefore, applying rigidity to kind terms leads to a revisionary account of nominal kind terms such as 'bachelor' and 'chair'. However, this seems to be an unfortunate consequence. Although we may be happy to grant that nominal kind terms cannot be given analytic definitions, it is a big step from this to conclude that they lack all descriptive content. A common strategy, among anti-descriptivists, is to dismiss associated descriptions as 'mere reference fixers', and not figuring in the semantics of the terms, thereby indicating that all of these descriptions are disposable and may fail to hold of the objects in the extension. But it is difficult to see how the descriptions associated with 'chair', 'mud', and 'pencil' could be dismissed in this way. For instance, it is difficult to imagine a world in which there are chairs but where all of the descriptions normally associated with 'chair' fail to hold. In the case of these terms, something like the original cluster theory seems much more plausible.

This suggests an alternative strategy if one wishes to single out the natural kind terms semantically: reject the idea that there is an interesting notion of rigidity that applies in the case of kind terms and appeal, instead, to the idea that natural kind terms, unlike nominal kind terms, are non-descriptional. This is the strategy endorsed by Soames 2002.

3.2 Non-descriptionality

According to Soames, the most important semantic concept in Kripke is not rigidity but non-descriptionality. This, he suggests, is true in the case of names as well, and hence the suggested parallel between names and natural kind terms can be upheld even without the notion of rigidity (2002: 264). Although speakers do associate descriptive properties with natural kind terms, Soames argues, these fail to provide necessary and sufficient conditions for something to be a member of a kind at a world, and sometimes the associated properties are not even true of actual instances of the kind—as, for instance, when speakers took whales to be fish. Hence, he affirms, 'the extension of a natural kind term is not semantically determined to be the set of objects that satisfy, at that world, the descriptive characteristics we (actual-world) speakers associate with the predicate' (2002: 266).

Now, whether indeed natural kind terms are non-descriptional in this sense is too large an issue to be properly addressed here. However, let me express some misgivings about using non-descriptionality as the distinctive mark of natural kind terms.

First, there are terms that are both widely recognized to be natural kind terms and to have a descriptive content—namely theoretical terms, such as 'H_2O'. Indeed, Soames himself suggests that, unlike 'water', 'H_2O' is semantically complex and synonymous with the description *something molecules of which consists of two hydrogen atoms and one oxygen atom* (2002: 308). It therefore has to be argued that non-descriptionality is the distinctive mark of *some* natural kind terms, but not all. But how is this category of non-descriptional natural kind terms to be singled out? This seems to be an even harder task. After all, the distinction between theoretical and non-theoretical terms is not that clear cut. Moreover, it is no good to appeal to the idea that it is the simple natural kind terms that lack descriptive content, since not all theoretical terms are complex (for instance, 'lepton').

Second, the claim that non-descriptionality is the distinctive mark of natural kind terms causes trouble when combined with the *a posteriori* proposal, according to which a term is a natural kind term only if it succeeds in picking out a natural kind.[25] It follows, from this combination of ideas, that it cannot be known whether a purported natural kind term has a descriptive content or not, prior to scientific investigation. If the liquid called 'water' has a unified, underlying structure, then 'water' lacks all descriptive content; if the liquid does not have such a unified structure, 'water' functions much like 'bachelor' and has a descriptive content. This is a rather astounding conclusion. Not only is it difficult to understand how the descriptionality of a term could, in this way, depend on the physical makeup of the world. The conclusion also introduces a problematic gap between how we use a given term (for instance, in reasoning) and its actual semantics.[26]

Third, there is a question concerning the motivations behind the claim that natural kind terms are non-descriptional. Although Kripke spends much effort arguing that names are non-descriptional, the arguments in the case of natural kind terms are much more swift and far less compelling. Take, for instance, Kripke's suggestion that just as all of the descriptions associated with 'Aristotle' may turn out to be false, so all of the descriptions associated with 'tiger' may turn out to be false. Since 'tiger' is used as a natural kind term, he suggests, we might find out that 'tigers had *none* of the properties by which we originally identified them. Perhaps *none* are quadrupedal, none tawny yellow, none carnivorous, and so on . . . ' (1980: 121). But it is very difficult to understand what we are supposed to imagine here—something that is a tiger but has none of the properties ordinarily attributed to tigers? Biologists would certainly have difficulties imagining this, in particular since they reject the microessentialist account of species.[27] Similarly, although there are (many) cases where particular beliefs about instances of a kind turn out to be false (as in the case of whales not being fish), this does not even begin to show that *all* of the associated descriptions are disposable.

Now, in the case of names, Kripke's modal arguments play a central role in the rejection of descriptivism. However, it should be clear that Kripke's claims about *a posteriori* necessity are less compelling in the case of natural kind terms. They carry conviction in the case of proper names, since statements such as 'Hesperus = Phosphorus' are identity statements and as such necessary, if true. But it is a serious question whether so called theoretical identity sentences are to be construed as proper identity statements. Most naturally, these sentences are construed as universally quantified conditionals (or bi-conditionals). What is required to get from such a statement to a proper identity statement is, first, that one appeals to the singular term use of these terms, rather than their predicative use. Second, we have to be given some reasons why the property identity statement should be accepted, and such reasons are not delivered by the acceptance of theoretical identity sentences within science.[28] What would have to be added are substantial, metaphysical assumptions about the identity of the *properties* involved (for instance, the assumption that H_2O = water)—assumptions that go well beyond what science establishes.[29] Hence, the parallel with names does not hold up. It cannot be argued that natural kind terms are non-descriptional simply on the grounds that we need a semantics of natural kind terms that secures the necessity of statements such as 'Water is H_2O'. It would also have to be shown that there is any necessity involved here in the first place (beyond that of nomological necessity).

I take all this to indicate that we should reexamine the widely shared assumption that natural kind terms are non-descriptional. Of course, there are proposals that attempt to fit descriptionality within the Kripkean framework, for instance, the reference-fixing proposal mentioned earlier, as well as various versions of two-dimensionalism. However, I think that the previous considerations indicate that we should look for a more radical alternative, one that involves a more comprehensive rejection of the Kripke-Putnam account of natural kind terms. If the arguments in the case of names do not carry over to the case of natural kind terms (when it comes to rigidity and necessity) and if there are reasons, both metaphysical and scientific, to question the Kripke-Putnam assumptions about the various kinds discussed (species, biological kinds, chemical kinds, etc.), it becomes unclear why we should remain fateful to the Kripke-Putnam framework in the case of natural kind terms.

Indeed, I take all this to suggest that Putnam was closer to the truth in his early discussions of law-cluster terms, before he introduced the talk of natural kind terms.[30] On this view, 'natural kind terms' do not have a different semantics than other kind terms—in both cases the cluster theory applies. The intuition that, even so, they differ from nominal kind terms is explained by appealing to what goes into the cluster of a natural kind term. Consider, again, chemical kinds. Paul Needham has suggested that the microessential conception of these kinds should be rejected, arguing that the microstructural level does not allow us to distinguish any nomologically interesting kinds—there is simply too much variation at this level. The question, he writes, is why the vast range of microscopic structures is associated with one and the same substance kind, rather than a genus of related substances:

> I suggest it is because macroscopic criteria determine sameness of substance kind, whose variable microstructure is then made the subject of scientific investigation . . . A macroscopically oriented account of sameness of kind doesn't challenge the claim that quantities of water have some appropriate range of microfeatures under specified conditions. But recognizing microproperties is not to favour them as more essential than others. If water is necessarily H_2O, it necessarily has its characteristic density too, characteristically reaching a local maximum at 4°C, it necessarily freezes at 0°C under normal atmospheric pressure, freezing at lower temperatures under higher pressures . . . , and so on for what science counts as water's essential macroscopic features. (2000: 21)[31]

This suggests that 'water' and other chemical kind terms are more like law-cluster terms than anything else. The meaning of these terms should be understood in terms of a set of properties, including properties relating to law-like behaviour, such that although not all of them can be rejected without a change in meaning, no single property (be it at the macro- or microlevel) can be singled out as the essential one.[32] Similar conclusions seem plausible when considering any of the specific kinds mentioned, such as biological kinds. For instance, species are not individuated by any underlying structural properties. Indeed, at this point there is little agreement on how to individuate species. As LaPorte notes, there are several competing conceptions of species, such as the biological species concept (which take interbreeding and reproductive isolation to be decisive) and the phylogenetic species concept (which appeals to ancestry and descent), which all divide the world into perfectly natural (but different) groups. Against this background, assumptions about the 'essence' or 'real nature' of species seem misguided. The picture that emerges, rather, is one that coheres with the cluster theory, where a set of interrelated properties (many of them easily observable) serves to delineate species.[33]

If I am right, therefore, we should reject the idea that there is a set among the kind terms, 'the natural kind terms', that are distinct from other kind terms at the level of semantic content. However, as noted in the introduction, it might be held that even if we cannot single out the natural kind terms as having a special *semantic content*, they can be singled out at the level of foundational semantics since their meaning is determined *externalistically*. The question is whether this could be defended once the previously mentioned scepticism concerning the first project is taken seriously. Are the difficulties surrounding the Kripke-Putnam semantics limited to the level of semantic content or do they spread to foundational externalism as well? To (briefly) consider this question, let us end by looking at one version of such a proposal, defended by LaPorte.

4. EXTERNALISM

As noted earlier, LaPorte denies that there is a difference between natural kind terms and nominal kind terms on the level of content. But he still holds that there is an important semantic distinction between natural kind terms and nominal kind terms: natural kind terms, unlike nominal kind terms, are linked to their referent (an abstract property) via causal contact with individual samples of the kind. Thus, they have their meaning determined externalistically:

> 'Whale' was baptized by a (probably informal) dubbing act: 'the term is to refer to *that* kind of thing' (the dubber points to some whales). On the causal theory what makes a thing belong to the extension of 'whale' will not be properties like *having a fish-like appearance* . . . that a speaker associates with whales, but rather *underlying* properties and relations that guarantee sameness of kind (to paradigm samples) . . . This stands in contrast to a term like 'bachelor', which is not causally grounded in sample bachelors. (2000: 304)[34]

'Externalism', in this sense, stands in opposition to the idea that meaning is determined by the internal states of the speaker, such as associated descriptions. It is not denied that descriptions play an important role in the determination of meaning. Indeed, LaPorte argues that descriptions are necessary for the 'dubbing act' to work and hence cannot be disposed of entirely. Rather, the claim is that associated descriptions are not *sufficient* to determine meaning: what is required, in addition, is the contribution of some property of the samples in question.[35] This, of course, is the type of externalism originally defended by Putnam in his Twin Earth thought experiment. The meaning of 'water', according to Putnam, is determined through an ostensive definition, *this is water*, such that 'water' refers to whatever kind that stands in the relation 'same liquid' to the sample in question. Given that the liquid pointed to by Oscar is H_2O, whereas the liquid pointed to by Toscar is XYZ, it follows (according to Putnam) that the term 'water' has a different meaning in Oscar's language than in Toscar's language—despite the fact that they associate all the same descriptions with the term.

According to this proposal, then, natural kind terms are distinct from nominal kind terms in how they receive their meaning—through a dubbing act—not in the semantic content expressed. Of course, nominal kind terms *could* be defined this way too. For example, it could be said that someone belongs to the kind 'swingle' if the person has the marital status of *these* people. In such a case, clearly, the external world contributes to determining the extension of 'swingle'. However, the idea is, in the case of natural kind terms, the external world provides a more distinctive contribution—what matters is *underlying* properties, essences that *cannot* be understood in terms of macrolevel features.

However, it should be clear, this brings us back to the difficulties discussed earlier. What is the underlying, essential property of whales such that something is a whale if and only if it has that property?[36] In the case of 'water', again, there is no underlying microstructural property of the required sort, and if Needham is right, individuating chemical kinds requires appealing to macrolevel features. Indeed, as noted earlier, LaPorte himself voices scepticism concerning this type of essentialism and stresses the hidden vagueness or 'open texture' of kind concepts. For instance, he argues, our pre-scientific use of kind terms is not such that it is decided that the term does not apply to various possible twin-substances. Rather, when encountering such substances, a *decision* is called for, a decision that, in effect, leads to a sharpening of the meaning of the term. This, LaPorte suggests, is illustrated by the historical case of 'jade'. In this case, speakers did not respond in the way predicted by Putnam since it was decided that both jadeite and nephrite were in the extension of the term. However, he stresses, we might as well have gone the other way; nothing about our earlier use with the term dictated an answer to the question whether jadeite was in the extension of 'jade'. Indeed, in the West ,early dictionary entries suggest that 'jade' only applies to nephrite, although this recommendation was not in fact followed. This, LaPorte suggests, teaches us something about Twin Earth. If we encounter a liquid that has all the superficial properties of our water, but a different underlying chemical composition, we have a *vague* case: 'We might call XYZ "water", contrary to Putnam. Then again, we might not: We could go either way' (2004: 100).[37]

However, if this is taken seriously, the question is what remains of the externalism that LaPorte appeals to. After all, to claim that we *could* go either way in Twin Earth scenarios is precisely to deny that the external feature plays a meaning-determining role: When the associated descriptions are the same, as in the case of Oscar and Toscar, the external feature will be decisive, or else its role is null. Even if we had gone the other way, externalism would fail to be supported—to support externalism it would have to be held that there was *no room for a decision*, that the underlying (external) essences were decisive. Scepticism concerning the existence of such underlying essences, therefore, leads to scepticism concerning the original externalist project.

What, then, about the widely shared Twin Earth intuitions? Do they not show that some form of externalism must hold for this class of kind terms? As always when it comes to intuitional evidence we have to be careful. First, there is the empirical question of how widespread these intuitions are (in particular outside the philosophical seminar room). As LaPorte notes, the case of 'jade' is as close as we get to a real-life Twin experiment, and in this case the evidence went the other way, failing to support the externalist conclusions.[38] Second, there is the much more difficult, general question, how we are to judge intuitions elicited by far-flung thought experiments. Twin Earth scenarios, as often noted, are nomologically impossible.[39] This should make us cautious since it means that we are supposed to have intuitions about cases that are far removed from the actual world. For instance, it may well be that our ordinary notion of a natural kind depends on the fact that there are no twin-substances, that observable and unobservable properties are nomologically related. If so, there are reasons to be sceptical of intuitions elicited by scenarios where this fact is assumed not to hold.

5. CONCLUSION

I have argued that we should question the widely shared assumption that there is a set of terms, 'the natural kind terms', that form a distinct semantic category among the kind terms. What, then, about the intuition that there is something special about terms that track natural, as opposed to nominal, kinds? Should it just be dismissed as mistaken? Well, even if these terms are not special in the sense that they form a special semantic category, they may be special in other senses—most obviously, they are special from the point of view of science. Although there is much disagreement as to how we are to characterize natural kinds, they do emerge as being of interest from an explanatory point of view: natural kinds support inductive generalizations. LaPorte, for instance, suggests that natural kinds are kinds with a certain explanatory value: 'A lot is explained by an object's being a polar bear. That it is a polar bear explains why it raises cubs as it does, or why it has extremely dense fur, or why it swims long distances through icy water in search of ice floes' (2004: 19).[40] In this respect, LaPorte suggests, natural kinds are different from highly unnatural kinds, such as the *named-on-a-Tuesday* kind, or kinds like toothpaste and trash.

Similar remarks, arguably, apply to the level of terms: 'polar bear' and 'water', unlike 'trash' and 'named-on-a-Tuesday', are natural kind terms in the sense that they have a special value from the point of view of prediction and explanation. If I say *Bob is a polar bear,* predictions can be made about Bob's behaviour and food preferences, whereas if I say *Bob was named on a Tuesday* the predictions that can be deduced are few and trivial. Hence, these terms are of special value from the point of view of science. This also means, plausibly, that they are of special value in ordinary contexts where projectible properties matter. (That Bob is a polar bear is a very good reason to stay away from Bob.) If so, it is fair to say that 'natural kind terms' *do* play a special role, only not a special *semantic* role.

NOTES

1. Goosens, for instance, writes that natural kind terms 'form a distinctive semantic kind' (1977: 149).
2. See, for instance, Haukioja (2006), Glüer & Pagin (forthcoming), LaPorte (2004), Schwartz (2002), and Soames (2002).
3. Bach is an early sceptic, suggesting that 'Putnam is implicitly making the undefended assumption that because a term applies to a natural kind, it belongs to a special semantic category, the category of natural-kind terms' (1987: 290).
4. I discuss this further in Wikforss 2008.
5. See, for instance, Donnellan 1993, Haukioja 2006, and McKinsey 1987. Another version of the *a priori* proposal can be found in Glüer & Pagin, forthcoming.
6. Haukioja, for instance, suggests that what characterizes natural kind terms is that although their normal application is based on *manifest* properties, 'they all possess criteria of correct application having to do with, for example, genetics or microphysical constitution, which are *non*-manifest and stable across worlds' (2006: 163).
7. For examples of this sort, see, for instance, Ben-Yami 2001 and Wilkerson 1993.
8. As several people have pointed out, there is in fact no unifying *microstructural* property of water or chemical kinds in general (see Needham 2000).
9. Naturally, there are responses available on part of the *a priori* proposal. Most commonly, it is suggested that the macrolevel properties can be utilised—either by appealing to the idea that in bad scenarios the term retains its character, and hence its meaning (although a full content will not be expressed) or by suggesting that there are two intensions associated with natural kind terms, one relating to the macrolevel properties (see Glüer & Pagin, forthcoming). All such accounts, though, have to grant that the term in question (construed as a singular term) fails to have a reference in such a scenario.
10. Jessica Brown suggests that even in the pre-scientific community 'it was part of the meaning of "gold" that it applies only to stuff with a particular hidden internal structure' (1998: 277). This seems rather doubtful, but even if it is granted it is unclear how the pre-scientific notion of a 'hidden internal structure' could the work required of picking out a unique underlying kind.
11. After all, even philosophers writing on the topic of natural kind terms have been unable to distinguish microstructure from chemical composition, and many have suggested that the essential property of species is microstructural (as assumption rejected by contemporary biology).
12. See, for instance, Brown 2004, Gallois 1996, McLaughlin and Tye 1998, and Korman 2006.
13. For a discussion of these theories, see, for instance, Griffiths 1997 and Häggqvist 2005.
14. As De Sousa (1984) stresses, it would also have to be the case that there are *several* natural kinds, rather than just one or two, if the claim that there are natural kinds (and natural kind terms) is to be of any interest.
15. The consequences of this idea are discussed in some detail in Häggqvist and Wikforss 2007.

16. See, for instance, Putnam 1970: 141 and Schwartz 1978.
17. See, for instance, Linsky 1984, LaPorte 2000, and Macbeth 1995.
18. For a discussion, see, for instance, Salmon 2005 and Soames 2006.
19. See, for instance, LaPorte 2000; López de Sa 2008; Martí 2004; Salmon 2003, 2005.
20. LaPorte (2000): 296 and 2004.
21. Martí (2004). See also Linsky (1984) and (2006) where he stresses the difference between rigid kind designators (such as 'blue') and non-rigid descriptions such as 'the colour of a cloudless sky at noon'.
22. López de Sa (2008) labels the second problem the 'over-generalization problem'.
23. See also Martí and Martínez-Fernández (this volume).
24. Recently, Soames himself seems to have given up on this requirement (2006 and forthcoming). Soames now accepts the notion of rigidity defended by Martí and others, and grants that it applies to simple non-natural kind terms as well, such as 'bachelor'.
25. And these two ideas are typically combined: Those who defend the *a posteriori* proposal tend to take natural kind terms to be wholly non-descriptional, whereas those who defend the *a priori* proposal take the meaning of natural kind terms to have a descriptional component (as on the reference-fixing proposal, mentioned earlier).
26. See Häggqvist and Wikforss 2007 for further critical discussion.
27. See LaPorte 2004. I return to the question of the individuation of species briefly following.
28. Soames recognizes this: ' "For all x, x is P iff x is Q" is necessary, if true, on the assumption that the identity statement involving the corresponding singular terms (*the kind p*, *the substance p*, etc.) is true: $tp = tq$' (2002: 260). He then goes on to argue that the identity in question can be delivered on purely semantic grounds. However, as often noted in the literature, this is very questionable since it seems clear that metaphysical assumptions about the identity of the relevant properties must be made. See Martí 2004 and Salmon 2003.
29. This is stressed by De Sousa 1984, Needham 2000, and Steward 1990.
30. This is discussed further in Häggqvist & Wikforss (unpublished).
31. See also Needham (forthcoming) and Sabbarton-Leary (unpublished).
32. Needham (forthcoming) also suggests that we should take seriously the idea that chemical kind terms can be given an account along the lines of the cluster theory, where stereotypical properties of the sort appealed to by Putnam, together with the more precise macroscopic properties appealed to by scientists, go into the cluster.
33. See also Dupré 1981. It should be noted that even on a cladistic classification, which is based on genealogy, macroscopic features play an essential role. For instance, as LaPorte stresses, there are intermediaries on the genealogical tree and decisions concerning these do turn on macroscopic similarities among the animals. There is also the question of how to delimit the beginning of a clade, and here systematists will appeal to various macrolevel features: such as the presence of feathers or the capacity of flight (LaPorte 2004: 84–5).
34. Schwartz concurs: 'It is the causal theory that explains how a term like "whale" is different from a term like "bachelor" . . . ' (2002: 274). Kripke, of course, also suggests that this is one respect in which natural kind terms are similar to names: in both cases, reference is fixed by an initial baptism (1980: 135). See also Soames 2002: 265.
35. It should be stressed that this version of the reference-fixing proposal is distinct from that discussed earlier, in the context of the *a priori* conception of natural kind terms, since it is *not* a view about the meaning of natural kind

terms but of the *determination* of meaning. Thus, it is possible to endorse this view of the determination of meaning and combine it with a wholly non-descriptional semantics (as Kripke does). On such a view the role of the reference-fixing descriptions is purely metasemantic.

36. Responding to Putnam's account, Dupré writes: 'My fundamental objection to the theory as a theory of biological kinds is that no such sameness relations suitable for Putnam's theory can be found in it' (1981: 70).
37. This claim, no doubt, is controversial and merits further discussion. However, it is worth noting that proper twin cases are nomologically impossible. That the meaning of our terms do not dictate what to say in nomologically impossible cases seems both plausible and, arguably, inconsequential—at least as far as communication goes.
38. To settle the empirical question, clearly, psychological experiments are required. Some have been undertaken in the case of natural kind terms. Jylkkä, Railo, and Haukioja (2009) have conducted experiments in Finland that they suggest provide evidence for externalism. Another experiment has recently been conducted by Machery and Olivola (2009). The study (so far unpublished) collects a large range of data from ordinary speakers in Mongolia, India, France, and the USA.
39. For a recent discussion of the difficulties involved in evaluating thought experiments, see Häggqvist 2009.
40. This view of natural kinds, of course, goes back to Quine 1969. For a discussion, see Häggqvist 2005 and De Sousa 1984. As De Sousa notes, on this view there will be no sharp boundary between natural kinds and other kinds and this is one reason it fails to underwrite the standard, Kripke-Putnam, conception of natural kinds.

REFERENCES

Bach, K. (1987) *Thought and Reference*. Oxford: Oxford University Press.

Ben-Yami, H. (2001) 'The semantics of kind terms', *Philosophical Studies* 102 (2): 155–84.

Brown, J. (1998) 'Natural kind terms and recognitional capacities', *Mind* 107: 275–303.

———. (2004) *Anti-Individualism and Knowledge*. Cambridge, MA: MIT Press.

De Sousa, R. (1984) 'The natural shiftiness of natural kinds', *Canadian Journal of Philosophy* 14: 565–80.

Donnellan, K. (1993) 'There is a word for that kind of thing: an investigation of two thought experiments', in J. E. Tomberlin (ed.), *Philosophical Perspectives, 7: Language and Logic*, Atascadero, CA: Ridgeview.

Dupré, J. (1981) 'Natural kinds and biological taxa', *Philosophical Review*, 91: 66–90.

Gallois, A. (1996) *The World Without, the Mind Within: An Essay on First-Person Authority*. Cambridge: Cambridge University Press.

Glüer, K. and Pagin, P. (forthcoming) 'General terms and relational modality', *Nous*.

Goosens, W. K. (1977) 'Underlying trait terms', in S. P. Schwartz (ed.), *Naming, Necessity, and Natural Kinds*, Ithaca, NY: Cornell University Press, 133–54.

Griffiths, P. E. (1997) *What Emotions Really Are*. Chicago: University of Chicago Press.

Häggqvist, S. (2005) 'Kinds, projectibility and explanation', *Croatian Journal of Philosophy*, 13: 71–88.

———. (2009) 'A model for thought experiments', *Canadian Journal of Philosophy*, 39: 55–76.

Häggqvist, S. and Wikforss Å. (2007) 'Externalism and *a posteriori* semantics', *Erkenntnis*, 67: 373–86.

———. (unpublished) 'Natural kind terms: resurrecting the cluster theory'.

Haukioja, J. (2006) 'Proto-rigidity', *Synthese*, 150: 155–69.

Jylkkä, J., Railo, H. and Haukioja, J. (2009) 'Psychological essentialism and semantic externalism: evidence for externalism in lay speakers' language use', *Philosophical Psychology*, 22: 37–60.

Korman, D. Z. (2006) 'What externalists should say about Dry Earth', *Journal of Philosophy*, 102: 503–20.

LaPorte, J. (1996) 'Chemical kind term reference and the discovery of essence', *Noûs*, 30: 112–32.

———. (2000) 'Rigidity and kind', *Philosophical Studies*, 97: 293–316.

———. (2004) *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press.

Linsky, B. (1984) 'General terms as designators', *Pacific Philosophical Quarterly*, 65: 259–76.

———. (2006) 'General terms as rigid designators', *Philosophical Studies*, 128: 655–67.

López de Sa, D. (2008) 'The over-generalization problem: predicates rigidly signifying the "unnatural" ', *Synthese*, 163: 263–72.

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

Macbeth, D. (1995) 'Names, natural kind terms, and rigid designation', *Philosophical Studies*, 79: 259–81.

Machery, E. and Olivola, C. (2009) Unpublished data.

Martí, G. (2004) 'Rigidity and general terms', *Proceedings of the Aristotelian Society*, 104: 129–46.

McKinsey, M. (1987) 'Apriorism in the philosophy of language', *Philosophical Studies*, 52: 1–32.

McLaughlin, B. and Tye, M. (1998) 'Externalism, Twin Earth, and self-knowledge',

in C. Wright, B. C. Smith, and C. Macdonald (eds), *Knowing Our Own Minds*, Oxford: Oxford University Press, 285–320.

Needham, P. (2000) 'What is water?', *Analysis*, 60: 13–21.

———. (2002) 'The discovery that water is H_2O', *International Studies in the Philosophy of Science*, 3: 205–26.

———. (*forthcoming*) 'Microessentialism: what is the argument?', *Nous*.

Putnam, H. (1962) 'The analytic and the synthetic', in H. Feigl and G. Maxwell (eds), *Minnesota Studies in the Philosophy of Science*, Minneapolis: University of Minnesota Press. Reprinted in Putnam 1975b, 33–69.

———. (1970) 'Is semantics possible?', *Metaphilosophy*, 1: 189–201. Reprinted in Putnam 1975b, 139–52.

———. (1975a) 'The meaning of "meaning" ', in K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science*, vol. 2, Minneapolis: University of Minnesota Press. Reprinted in Putnam 1975b, 215–71.

———. (1975b). *Mind, Language and Reality, Philosophical Papers*, vol. 2, Cambridge, MA: Cambridge University Press.

Quine, W. V. (1969) 'Natural Kinds', in N. Rescher (ed.), *Essays in Honor of Carl G. Hempel*, Dordrecht, Netherlands: D. Reidel, 5–23.

Sabbarton-Leary, N. (unpublished) 'Isomerism and isomorphism: some problems for metaphysical necessity'.

Salmon, N. (2003) 'Naming, necessity, and beyond', *Mind*, 112: 475–92.

———. (2005) 'Are general terms rigid?', *Linguistics and Philosophy*, 28: 117–34.

Schwartz, S. P. (1978) 'Introduction', in S. P. Schwartz (ed.), *Naming, Necessity, and Natural Kinds*, Ithaca, NY: Cornell University Press.

———. (2002) 'Kinds, general terms, and rigidity: a reply to LaPorte', *Philosophical Studies*, 109: 265–77.

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

———. (2006) 'Reply to critics', *Philosophical Studies*, 128: 711–38.

———. (forthcoming) 'What are natural kinds?', *Philosophical Topics*.

Steward, H. (1990) 'Identity statements and the necessary *a posteriori*', *The Journal of Philosophy*, 87: 385–98.

Wikforss, Å. (2005) 'Naming natural kinds', *Synthese*, 145: 65–87.

———. (2008) 'Semantic externalism and psychological externalism', *Philosophy Compass*, 3: 158–81.

Wilkerson, T. E. (1993) 'Species, essences and the names of natural kinds', *Philosophical Quarterly*, 44: 1–19.

Wolf, M. P. (2002) 'The curious role of natural kind terms', *Pacific Philosophical Quarterly*, 83: 81–101.

5 The Commonalities between Proper Names and Natural Kind Terms

A Fregean Perspective

Harold Noonan

How can we explain the commonalities Kripke brings out in the third lecture of *Naming and Necessity* (*NN*, 1972) between proper names and natural kind terms? Since Kripke has, in his second lecture, already given what he regards as decisive arguments against the Frege-Russell 'description theory' of proper names, he sees his observations about natural kind terms as merely providing more grist to his mill—that is, his more-Millian-than-Mill rejection of the application of any notions of connotation, sense, or descriptive meaning, not only to proper names, but also to certain Millian 'general names'. However, for someone like me, unconvinced by Kripke's arguments, the question arises how these commonalities can be explained consistently with some form of the Frege-Russell account. Specifically, for me, the question is how to explain the features of natural kind terms that Kripke describes from a Fregean perspective.

In what follows, I sketch, in the course of responding to Kripke's modal argument against descriptivism, a Fregean account of proper names, including, in particular, a Fregean explanation of why they function as rigid designators. I defend this account against Kripke's argument from ignorance and error and briefly indicate how it provides explanations of some of his examples of the necessary *a posteriori* and the contingent *a priori*. I turn next to Kripke's story about natural kind terms. Here my aim is to show that everything acceptable in Kripke's story is consistent with Fregean theory and that the distinctive Kripkean claims—about the special character of natural kind terms, contrasted with other general terms, as rigid designators, the essences of natural kinds and the metaphysically necessary *a posteriori* status of theoretical identifications—can all, rightly understood, be accounted for in a Fregean setting.

1. THE MODAL ARGUMENT AND (FREGEAN) RIGID DESIGNATION

Kripke employs two types of argument against the description theory of proper names: the modal argument and the argument from ignorance and error.

The modal argument starts from the observation that proper names are rigid designators. The sentence:

(1) The inventor of bifocals might not have been the inventor of bifocals

is ambiguous. There is a scope ambiguity comparable to the scope ambiguity in Russell's example, 'George IV wondered whether Scott was the author of *Waverley*'.

By contrast (2), following, is unambiguous:

(2) Benjamin Franklin might not have been Benjamin Franklin.

This difference between (1) and (2), Kripke says, is explained by the fact that the name 'Benjamin Franklin' is a rigid designator, whose designation in any possible world is its *actual* designation, whereas 'the inventor of bifocals' is a flexible designator. So the name and the description are not synonymous. And, in general, names behave like 'Benjamin Franklin', that is, as rigid designators, whereas descriptions, or at least those that might at all plausibly be claimed by a description theorist to be synonymous with names, behave like 'the inventor of bifocals'; that is, they are flexible designators. So the description theory, as a theory of meaning anyway, is refuted.

What I wish to do now it to suggest how the rigidity of proper names might be explained within a Fregean framework.

Frege never explicitly discussed modal contexts. However, they are like propositional attitude contexts in blocking substitutivity *salva veritate* of codesignating singular terms (definite descriptions). Within a Fregean framework this failure of substitutivity can be accommodated only by regarding such codesignating singular descriptions as having indirect reference in modal contexts, i.e. as having reference to their customary senses. In general, a Fregean account of modal operators must treat them, like propositional attitude verbs, as creating contexts in which reference shifts occur.

Now the ambiguities present in such modal sentences as (1) are also present in ascriptions of propositional attitudes such as:

(3) George IV wondered whether the author of *Waverley* was a Scot.

From a Fregean viewpoint the ambiguity in the latter has to be explained in something like the following way. On one reading of (3) it asserts that a relation (of wondering whether) holds between George IV and the thought identified in the that-clause, the thought that the author of *Waverley* was a Scot. On this reading of (3), 'the author of *W*' has indirect reference after 'George IV wondered whether'. On the other reading of (3) 'the author of

W' retains its direct reference after 'George IV wondered whether'. The sentence therefore asserts *of* the author of *W* that George IV wondered whether he was a Scot, that is, that thinking of the author of *W* in some way, i.e. under some mode of presentation, George IV wondered whether the person so presented was a Scot.

Peter Geach (1976) (see also Kaplan [1969]) brings out the ambiguity as follows. He uses the term 'aspect' to mean the Fregean sense of a proper name, and speaks of an aspect α as an aspect 'of' an object x when, in Fregean terminology, α is a mode of presentation of x, a way of thinking of x. Next he stipulates that '[α is F]' is to stand for the thought composed of the aspect α and the sense of the predicate 'is F'. The thought that [α is F] is the thought you would express in language by attaching the predicate 'is F' to a subject term whose sense is the aspect α.

The two readings of (3) are now:

> (3*) George IV wondered whether [the author of *W* was a Scot]

and

> (3**) for some α, α is an aspect of the author of *W* and George IV wondered whether [α was a Scot].[1]

The relevant ambiguity in (1) can now be brought out by the two readings:

> (1*) It might have been the case that [the inventor of bifocals was not the inventor of bifocals]

and

> (1**) for some α, α is an aspect of the inventor of bifocals and it might have been the case that [α was not the inventor of bifocals].

In both cases some restriction on type of aspect is required if the second reading is to capture the intuitive English meaning.

In the case of propositional attitude ascriptions we do not invariably allow exportation of a singular term from an opaque position within a propositional attitude context to a transparent position outside it. To illustrate with the familiar examples (Quine 1966; Kaplan 1969): we do not allow the inference from 'Ralph believes that the shortest spy is a spy' to 'Ralph believes of the shortest spy that he is spy' and thence to 'there is someone Ralph believes to be a spy'. Nor do we allow the inference from 'Ralph knows that the shortest spy is a spy' to 'Ralph knows of the shortest spy that he is a spy' and thence to 'there is someone Ralph knows to be a spy'. Nor, finally, do we allow the inference from 'Ralph knows that the shortest spy is the shortest spy' to 'Ralph knows of the shortest spy that he is the shortest spy' and thence to 'there is someone Ralph knows to be the shortest spy' or 'Ralph knows who the shortest spy is'.

Similarly, from the fact that George IV wondered whether the author of *W* was a Scot it is not immediately evident that we should infer that there was someone *of* whom he wondered whether he was a Scot.

To accommodate these examples, the Fregean must say that the existential form (3**) captures the intuitive English meaning of (3) with the description given a wide scope reading only if the quantification is restricted to what we might call 'identity-revealing' aspects—where an aspect α is identity-revealing only if knowing that [α is *X*] suffices for knowing who is *X*.

An identity-revealing aspect may be what is expressed by what Kaplan (1969) calls a 'vivid name'. Alternatively, of course, the notion of an identity-revealing aspect may well be context dependent. What counts as knowing who x is may vary from context to context. But in George IV's context he did not know who the author of *W* was, although he knew that the author of *W* was the author of *W*, because there was no identity-revealing aspect α of the author of *W* such that he knew that [α was the author of *W*].

Similarly, in order to capture the intuitive English meaning of (1), the modal statement about the inventor of bifocals, on the reading on which the first occurrence of the description is taken as lying outside the modal operator, we need to regard the quantification in (1**) as restricted to what we might call 'essence-revealing' aspects. Suppose that Benjamin Franklin was essentially a human being, so 'the inventor of bifocals might not have been human', on the intended wide scope reading, is false. Still, bifocals might not have been a human invention—like Velcro, according to *Star Trek* mythology, they might have been invented by Vulcans. So:

> for some α, α is an aspect of the inventor of bifocals and it might have been that [α was not human]

is true if the quantification ranges over all aspects, since 'the inventor of bifocals' expresses an aspect of Benjamin Franklin and it might have been that [the inventor of bifocals was not human].

So we need the notion of an essence-revealing aspect of an object. An aspect α is an essence-revealing aspect of an object *X* iff it is necessary that [if anything is α it is *F*] for any '*F*' such that '*F*' expresses a necessary property of *X*, and for no '*F*' such that '*F*' expresses an accidental property of *X* is it necessary that [if anything is α it is *F*]. An essence-revealing aspect of an individual, in other words, is one that presents it as the possessor of *all and only* its essential properties. (But what of individuals that do not have individual essences, but only general essences? Suppose, for example, that I am essentially a human being but have no other essential property [so that I could have been anything that any human being could be]. And what of counterexamples to the Identity of Indiscernibles, such as Max Black's two spheres? The definition of an essence-revealing aspect does not need to be modified to deal with these possibilities, but the notion of an aspect needs generalization. We need to define an aspect of an individual *X* as the sense of a predicate satisfied by *X*. Aspects in the original sense defined, the senses of Fregean proper names can now be thought of as the senses of predicates of the form 'is identical with *X*' where '*X*' is the Fregean proper name. So the aspect expressed by 'Socrates' can now be thought of as the sense of the predicate 'is identical with Socrates' or Quine's 'Socratizes'. The treatment of quantification into propositional attitude contexts now needs to be understood as restricted to aspects as originally defined.)

Of course, it may be that this notion of an essence-revealing aspect, like that of an identity-revealing aspect, will have different applications in different contexts. Maybe we have no context-independent notion of an essential property. And, of course, I have not explained the notions of an essential property and an accidental property by appeal to the notion of an essence-revealing aspect—I have simply taken them for granted in introducing the latter.

Now we are in a position to give a Fregean account of what the rigidity of proper names comes to. Namely, that there is a convention in force whereby a proper name, as opposed to a description like 'the inventor of bifocals', must not be used in a modal context to refer to its indirect reference, i.e. its customary sense, and in general, must not be used by a speaker to refer to the sense *he* associates with it in contexts where replacement by a proper name with the same direct reference but a different customary sense is not guaranteed to preserve truth-value. With this convention in force there is no reading of (2) corresponding to the reading of (1) as (1*). It will be obvious that this proposal is similar to Dummett's proposal that names (that have senses that may be expressed by descriptions) may be regarded as abbreviations for descriptions conventionally required to take wide scope in modal contexts. In fact, my proposal is not an alternative to Dummett's; it *is* Dummett's in Fregean guise. Dummett would agree, since on his view quantification into contexts of alethic modality, like quantification into any intensional context, is to be understood in terms of quantification over senses (Dummett 1981: Ch. 9).

But what is the rationale for this convention? Frege thinks that ordinary proper names vary in sense from speaker to speaker. But we aim in conversing with others to speak about the same things as they do. If we fail to do this, if we lack a common subject matter, we fail to communicate at all. But if the senses of proper names vary from speaker to speaker in the way Frege thinks, we *will* lack a common subject matter if we use them to refer to their indirect referents, i.e. their customary senses, i.e. the senses *we* associate with them. For then if I make an assertion containing a name and you repeat it, you may speak falsely though I spoke truly.

If the convention has this rationale, other behaviours not involving the use of proper names to *refer* to senses will be ruled out for similar reasons. If I say, 'Cicero is Tully. That is necessarily true', what I say may be true if the thought I express with the sentence 'Cicero is Tully' is a necessarily true thought because I associate the same sense with the two names and I refer to that thought when I utter 'That' in my comment. But if you associate different senses with the names from me (and different senses with the two names), what you say, if you use 'That' to refer to the thought you express with 'Cicero is Tully', may be false. So if you repeat what I say but use 'That' in your comment to refer to the thought you expressed with your previous sentence and I use 'That' in my comment to refer to the thought I expressed in my previous sentence, there is again no guarantee that our utterances will have the same truth value, since they may have different subject matters. This indicates how a Fregean should respond to the reformulation Kripke gives of his modal argument in the preface to the book version of *NN*, where he argues that Dummett's suggested scope convention cannot account for rigidity.

The crucial point can be made independently of the Fregean apparatus, of course. Lewis, citing Noonan 1979, states as the first of the seven points that should be taken to heart by a descriptivist: 'there may or may not be rigidification. If there is, that will avoid confusion between people who have attached the same term to the same referent by means of different descriptions. For nothing will be true as one person means it but false as the other means it, not even when the term appears in modal contexts' (1984: 223). The closest I have seen to the account of rigidity given in the text in another writer[2] is in Bob Hale's 2004, where he writes: 'We could say: *d* is (used as) a rigid designator if and only if it (is used in such a way that it) refers when it figures inside counterfactual or other modal contexts to what it refers to, if to anything at all, outside such contexts' (2004: 362). He suggests that this might be truer to Kripke's intentions than the usual definition because it contains no commitment to an ontology of possible worlds and makes no use of any notion of transworld identity. He notes that it makes the rigidity of a proper name a matter of our intentions in using it and in a footnote comments that Kripke's thesis that proper names are rigid designators, so understood, goes flatly against Frege's view that words in an indirect context undergo a systematic shift of reference—at least if modal contexts are taken as indirect. Note this difference between Hale's suggestion and mine, however. He is putting forward an account of what rigidity *is*. My proposal is less general; it is intended merely as an explanation of why *proper names* function as rigid designators, but is not intended to apply to all rigid designators, and specifically not to rigid descriptions (descriptions whose rigidity is what Kripke calls 'de facto'). Another difference, of course, is that Hale does not explain *why* we should use proper names as rigid designators, or what purpose it serves.

Of course, this suggestion appeals to a feature of the ordinary use of proper names that Frege regards as a *defect*; namely, variation in sense from speaker to speaker, but we can see that this is not a fatal objection by turning now to Kripke's argument from ignorance and error.

2. PROPOSITIONAL ATTITUDES, DEFERENTIAL INTENTIONS, AND FREGE'S PUZZLE

I shall not go through the familiar examples and arguments in *NN*, but simply record the familiar response found in a host of writers, which seems to me to be correct. What Kripke has to show is that one can refer to a thing by a name even when one knows of no descriptive identification which determines the reference of the name, *not even* one which involves a first-person reference to oneself. But none of his examples show this. To illustrate: while, as Kripke says, one does not use the name 'Gödel' to refer to the unknown Schmidt just because Schmidt proved the incompleteness of arithmetic and the only biographical detail one associates with the name 'Gödel' is 'prover of the incompleteness of arithmetic', this proves nothing as strong as Kripke claims. For to use 'Gödel' to refer to Gödel in this circumstance one must know that there is just one person who is named 'Gödel' and from references to whom by that name one's familiarity with the name derives.

In general, such an egocentric, metalinguistic, causal description is bound to be available, if nothing else is, whenever one is capable of referring to something by a proper name.

This argument can be reinforced by thinking about the picture Kripke sketches of the way in which a name becomes established in a community. There is an initial baptism, in which the baptiser intends the reference of the name to be fixed by a description or ostensively (which Kripke suggests may be assimilated to the former case). The name is then passed on to a second user and so on down a chain. When the second user picks up the name he must have the intention to use it to refer to whatever was referred to by the person from whom he is picking it up; he cannot decide to initiate a new use of the name as a name for his pet aardvark, for example. And the same must be true of all subsequent users. Intention thus enters in a crucial role twice into Kripke's picture. Now when the second user picks up the name he will not count as intending to use the name with the same reference as those from whom he acquired it if he does in fact intend to use it as a name for his pet aardvark and, in fact, quite unbeknownst to him, the people he heard using the name *were* using it as a name for his pet aardvark. Rather, it must be that the content of his intention is: to use the name to refer to whatever was the referent in the mouths of those from whom I heard it. Similarly, when the baptiser introduces the name he must have an intention whose content is: to use the name to speak of the so-and-so (or *that* object if he fixes the reference of the name ostensively). But then neither the second user's first uses of the name, governed by his deferential intention, nor the baptiser's first uses, governed by his baptismal intention, can possibly be counterexamples to the Frege-Russell description theory as Kripke defines it.

Counterexamples can only emerge once the second user (to focus on him) *ceases* to have the intention to refer to whatever the people from whom he acquired the name used it to refer to *and* acquires no other reference-determining intention in its place (he cannot acquire the intention to refer to whatever he most recently referred to with the name, for example, or to what others in his community now refer to by the name, or to Kripke's or his own pet aardvark). Similarly, uses of the name by the baptiser can only emerge as counterexamples to the description theory once the baptiser *ceases* to have the intention which originally fixed the reference of the name, *and* acquires no other reference-determining intention in its place, not even the intention to refer to whatever he previously used the name to refer to. Such uses of the name can be counterexamples only if their reference is the object standing in a certain causal-historical relation *R* to them, to be described by Kripke—which requires that an object *X would be acknowledged*[3] by the name user to be the reference if evidence were provided to him that *X* did indeed stand in relation *R* to his utterances—but the name user has no *intention* to refer to the object standing in relation *R* to his utterances, nor any other intention capable of determining that object as his reference. But Kripke never describes a case satisfying these requirements. In fact, the intelligibility of the stories he tells depends crucially on the fact that the user of 'Gödel', or whatever, is a member of a community and intends to use the name in the same way as others—it is only this that makes it plausible that, despite his ignorance and mistaken beliefs, he does not refer to *Schmidt* but does refer to *someone* when he utters the name.

The picture of our naming practices now developed and suggested (for one looking at the phenomena through Fregean eyes, at least) by Kripke's attempted counterexamples to the description theory is one according to which some users of a proper name are *deferential* in their use of it to others and need have no knowledge of the name's bearer which can be expressed without reference to those others and to their uses of the name. But within a Fregean framework this is a version of the view that the sense of a proper name varies from speaker to speaker and the correctness of that view, as we saw, provides a rationale for the convention proposed earlier as an explanation of the fact that proper names function as rigid designators. We do not need to think of variation in sense, understood in this way, as a defect, endangering the unity of the linguistic community, so long as such variation is accompanied by an acknowledgment by members of the community of their responsibility to achieve a common reference.

I conclude that Kripke's attack on the Fregean position fails to refute it. But it teaches us the importance of the social dimension of language and the crucial role of deferential intentions in determining name reference. However, this can be accommodated within a Fregean framework and must be added to it if we are to go beyond Frege's dismissal of ordinary language and describe it correctly.

But what of the role of proper names within propositional attitude contexts, and Frege's puzzle of identity (to solve which Frege made the distinction between direct and indirect reference)?

There is intuitively a difference between George IV's believing *of* Scott that he is a Scot and his believing *that* Scott is a Scot, just as there is between his believing of the author of *W* that he is a Scot and his believing that the author of *W* is a Scot. 'George IV believed that Scott was a Scot' (like 'George IV believed that the author of *W* was a Scot') is ambiguous. It can mean (a) for some α, α is an (identity revealing) aspect of Scott and George IV believed that [α was a Scot], or (b) for some α, α corresponds to the sense of 'Scott' in my idiolect and George IV believed that [α was a Scot]. Here correspondence is that relation between senses which is required if an ascriber is to be correct in his ascription, via the use of a proper name, of a propositional attitude to an ascribee. Correspondence is not identity. If two people use a name to refer to the same object, one deferentially to the other, the senses they associate with the name are different, but correspondent. Kripke's 'Pierre', before his kidnap, uses 'Londres' in a sense which is different from, but corresponds to the sense I associate with 'London', which is why we say that Pierre believes that London is pretty (Kripke 1979). Correspondence does, however, require identity of reference (though both senses may have no reference). But identity of reference is not sufficient. It may be that nothing more general can be said and that what correspondence requires over and above identity of reference is contextually determined (see for further discussion Noonan 1979 and 1981; there are other similar discussions in the literature; the closest I have seen is Chalmers [unpublished]).

The solution to the puzzle of identity is now evident.

> *X* believes that Hesperus . . .

and

> *X* believes that Phosphorus . . .

differ in truth-conditions because the first has the truth-condition:

> for some α, α corresponds to the sense I express with 'Hesperus' and *X* believes that [α . . .]

whilst the second has the truth condition expressed by replacing ' "Hesperus" ' with ' "Phosphorus" '.[4]

We can now turn to Kripke's examples of the necessary *a posteriori* and contingent *a priori*.

Kripke cites as examples of the necessary *a posteriori* identity statements flanked by two proper names, e.g. 'Cicero is Tully'. It follows from the account sketched earlier that in most people's mouths this will express a contingently true thought, different thoughts in different mouths, of course. The reason for the appearance of necessity is that when the operator

'Necessarily' is prefixed the resultant statement is one to which the convention that one must not use a proper name to refer to the sense one associates with it applies. Hence, 'Necessarily, Cicero is Tully' must be understood not as ascribing a property to the thought that a free-standing occurrence of 'Cicero = Tully' would express, but as saying of the objects Cicero and Tully that they are necessarily identical, and since an identity pair is essentially an identity pair, this is true. But there is no *thought* here that is both necessarily true and knowable only *a posteriori*.

Similarly, if we may take it that it is an essential property of any human being that he is composed of molecules, the Kripkean story involves that it is a necessary *a posteriori* truth that Kripke is composed of molecules (if he exists). But the Fregean account is different. 'Kripke is made up of molecules' will express in most people's mouths a contingently true thought, different ones in different mouths. But if I prefix the operator 'Necessarily' I cannot be understood as ascribing a property to the thought that I express with that sentence. Rather, I have to be understood as saying of the direct reference of the name in my mouth, i.e. Saul Kripke, that he is necessarily made up of molecules. And given that being made up of molecules is an essential property of Kripke—which is something the Fregean need not deny[5]—what I say will be true. But nothing will be both necessarily true and knowable only *a posteriori*. (According to the account I have proposed if Kripke is essentially made of molecules there will be a necessarily true thought, the thought that [if something is X it is made of molecules] where 'X' expresses an essence-revealing aspect of Kripke. But this thought need not be knowable only *a posteriori*. Consider another example. I can fix the reference of a name [rigid designator] 'NP' by the description 'the number of planets'. Having done so I can know, but only *a posteriori,* that NP is odd. I can also know *a priori* that if NP is odd, NP is essentially odd. So I can infer that NP is essentially odd, which I can only know *a posteriori*. Hence I can know, but only *a posteriori*, that I will speak the truth if I say 'Necessarily, NP is odd'. But it does not follow that there is any necessary *a posteriori* truth expressible in the form 'if something is *X* it is odd'. The necessary truth 'if something is *X* it is odd' such that '*X*' expresses an essence-revealing aspect of the number of planets can be: 'if something is 9 it is odd'.)

From a Fregean perspective, then, the Kripkean notion of a metaphysically necessary *a posteriori* truth is a conflation of two notions, that of an essential property of an object and that of an only *a posteriori* knowable contingently true thought. What Kripke identifies as a statement which is (or expresses) a necessary *a posteriori* truth is merely one expressing an only *a posteriori* knowable thought in which an object picked out by a proper name, i.e. by an expression subject to the convention discussed, is ascribed an essential property. Kripke asserts (1971: 180) 'the notion of essential properties can be maintained only by distinguishing between the notions of *a priori* and necessary truth'. That is what I am denying.

These comments on the necessary *a posteriori* carry over, *mutatis mutandis*, to the contingent *a priori*. Kripke gives the example 'the length of this rod is one metre'. Another example acceptable to him would be Dummett's 'St. Anne was a parent'. In the mouth of someone who uses 'St. Anne' with the sense of 'the mother of the Blessed Virgin Mary', this expresses an *a priori* knowable, necessarily true thought (or rather the conditional 'If St. Anne existed . . . ' does). The reason for the appearance of contingency is that when the operator 'Necessarily' is prefixed the result is a statement to which the convention applies that one must not use a name to refer to the sense one associates with it. Hence one can only be understood as speaking, not of the sense one associates with 'St. Anne', but of the person herself. But since she was only contingently a parent, one's statement has then got to be understood as saying something false. But there is no thought here which is both contingently true and knowable *a priori* which one expresses when one says 'St Anne (if she existed) was a parent'.

We now have the wherewithal to discuss Kripke's contentions about natural kind terms.

3. TWIN EARTH AND FREGEAN NATURAL KIND TERMS

The initial focus of Kripke's argument is that natural kind terms cannot be defined by a list of superficially observable properties, such as can be found listed for 'tiger' in the dictionary. Nor can an appeal to a weighted majority or cluster of these properties serve to define the kind. We might find out that tigers have none of the properties listed. Even if this is not so, there might be tigers elsewhere that have none of them. And even if this is not so, there might have been tigers that had none. Conversely, there might have been, and for all we know might be, creatures with all the properties by which we identify tigers that are not tigers because reptilian. Similarly gold might not be yellow. It might not have any of the superficial properties by which we recognize it. Conversely, there might have been, for all we know there might be, and, in fact, there actually is, something with all the superficially observable properties of gold that is not gold—fool's gold.

Kripke's discussion here is always linked in the literature to Putnam's discussion of his Twin Earth examples, which can be seen as illustrating some of the same points. The substance XYZ on distant Twin Earth is no more water than the reptilian fool's tigers are tigers. It is just indistinguishable from water by superficially observable properties.

The examples and arguments are well known. But the point I wish to make should be obvious from the preceding. Nothing at all in Kripke's and Putnam's arguments is in any way inimical to the Fregean viewpoint. The lesson to be learned is perhaps best expressed in Putnam's claim that the notion of indexicality extends beyond the obvious cases. Natural kind terms have an indexical element and natural kinds are not identified by a set of superficially observable properties, but by a set of samples and a relation something must have to the samples to be a member or instance of the kind—where the relation in question is not merely a matter of possessing certain superficial properties in common but, in the case of biological kinds, has something to do with internal structure, and in the case of kinds of substance, has something to do with chemical composition.

But this is entirely in conformity with the Fregean viewpoint. Kripke is right to emphasize the commonalities between proper names and natural kind terms, but these are not in any way an argument against a Fregean account of the latter. I may introduce a proper name for a person whom I encounter only in certain circumstances, where his most salient features are a certain set of observable properties. But I can certainly speculate, having done so, that he might in other times and places look different and behave quite differently, or anyway that he might have looked quite different and behaved quite differently. And I can also speculate that he does, in fact, look quite different and behave quite differently (I can speculate that he is in disguise and/or we are in fact trapped in a hall of distorting mirrors). And, of course, it is part of my understanding the name as a name of a person that I acknowledge that someone else, someone other than the person I have fixed as the reference of the name, might have had and indeed might have all his observable properties.

All of these platitudinous remarks are entirely compatible with the Fregean account previously sketched, and *mutatis mutandis* Kripke's observations are compatible with a Fregean account of natural kind terms.

Kripke presents his picture of how natural kind terms function in the following passage, emphasizing the likeness to the case of proper names:

> In the case of proper names, the reference can be fixed in various ways. In an initial baptism it is typically fixed by an ostension or a description. Otherwise, the reference is usually determined by a chain, passing the name from link to link. The same observations hold for such a general term as 'gold'. If we imagine a hypothetical (admittedly somewhat artificial) baptism of the substance, we must imagine it picked out by some such definition as 'Gold is the substance instantiated by the items over there, or at any rate, by almost all of them' . . . I believe that in general names for natural kinds (e.g., animal, vegetable and chemical kinds) get their reference fixed in this way; the substance is defined as the kind instantiated by (almost all of) the given sample. The 'almost all' qualification allows that some fool's gold may be present in the sample. If the original sample has a small number of deviant items, they will be rejected as not really gold. If on the other hand, the supposition that there is some uniform substance or kind in the initial sample proves more radically in error, reactions can vary; sometimes we can declare that there are two kinds of gold, sometimes we may drop the term "gold" . . . the original samples get augmented by the discovery of

> new items. . . . More important, the term may be passed from link to link exactly as in the case of proper names, so that many who have seen little or no gold can still use the term. Their reference is determined by a causal (historical) link. (*NN*: 328)

On the Kripkean picture, the term 'gold' acquired its reference in the way a proper name acquires its reference, and especially the way a proper name like 'Jack the Ripper' acquired its reference. 'Jack the Ripper' was initially fixed to have as its reference: the man who committed those murders, or at any rate most of them. If evidence had emerged that a few of the murders were committed by a different person (a copycat killer), the reference of the name would have remained the man who committed most (or maybe the most gory) of them; the victims of the copycat killer would not have been thought of as victims of Jack the Ripper. But reactions might have varied if honours, as it were, had turned out to be equally divided. Be that as it may, the name's reference was originally fixed as: *that man*, the man who committed *those* murders, whoever he is—and it was understood that it was the job of the experts (the detectives) to find out. So 'gold', on Kripke's picture, was originally introduced just as a name for: *that* kind of substance, the kind instantiated by *those* samples, whatever it is—and it was understood that its nature was a matter for future discovery. So it was entirely comprehensible to its first users that a substance absolutely indistinguishable from gold by all the tests available to *them* for being gold might not be gold, but a different kind of stuff.

Just because of the parallelism Kripke insists on, however, we can see that there can here be no conflict with the Fregean view. At the time of the initial baptism the users of the natural kind term 'gold' use it with the intention to refer to: *the kind of stuff of which these samples are instances.* Subsequently, users intend to refer to what the first users refer to—they use the term deferentially. Kripke, in fact, does not attempt to sketch out in the case of natural kind terms, as he does in the case of proper names, a counterexample to the Fregean view. Rather, he writes as if he has already done so in his account of how natural kind terms are introduced. But this account only shows that natural kind terms are not equivalent to descriptions free of names, demonstratives, and other indexicals, just as, as everyone agrees, proper names of persons are not equivalent to descriptions free of names, demonstratives, and other indexicals; it does not challenge the Fregean requirement of identifying knowledge for reference.

There are, of course, doubts about whether Kripke's account of the way natural kinds terms are secured a reference can be correct. The intuition that 'gold' or 'water' might have been introduced in a primitive community to name a kind that members of the community could not distinguish, given their limited knowledge, from a kind instantiated elsewhere (say, on a distant Twin Earth) is highly plausible.[6] But what seems most readily to make sense of it is the thought that the term could have been introduced to name the *narrowest*, *the most uniform*, kind instantiated by the samples, so that even if the stuff on Twin Earth were indistinguishable from that on Earth by all the means available to the community, they could still make sense of the idea of differences *they* could not discern which made the difference. And, indeed, it is only charitable to assume that some such restriction must be intended by Kripke because, of course, kinds form a hierarchy, both in the biological and the chemical case, so to talk of '*the* kind instantiated by the samples' is necessarily to fail to denote. If Kripke does intend such a restriction it fits well with Putnam's remarks about 'jade' (where the samples are assumed to be mixed) and with our intuitions (or those of most of us) about XYZ on Twin Earth. But it does not fit well with all our intuitions. (It is, of course, particularly difficult to see how the restriction can give the right results in the case of biological kinds since every individual organism is different, genetically and in its history, from any other. Yet the members of a primitive community could presumably, on Kripke's story, pick out the kind *tiger* by reference to a single tiger: 'the kind of creature of which *that* is an example', or by reference, say, to a female tiger and its cubs.) Gold naturally occurs in just one isotope, so any baptismal sample for 'gold' on Earth consists entirely of that isotope, but it seems that the term 'gold' could have been introduced in the way Kripke sketches by a primitive community without its extension being composed only of samples of the naturally occurring isotope, as the restriction would require. Similarly, if all our baptismal samples for 'water' had happened to be H_2O (light water), the term 'water' could not, counterintuitively, have been introduced in the way suggested, given the restriction, and at the same time have been given an extension including samples of heavy water. The restriction does not even accord with all of Kripke's intuitions. He remarks that if this substance (H_2O) 'can take another form—such as the polywater allegedly discovered in the Soviet Union with very different identifying marks from that of what we now call water—it is a form of water because it is the same substance' (*NN*: 323).

An alternative way of elaborating, or modifying, Kripke's proposal, which might give intuitively acceptable results, would be to propose that the introducers of a natural kind term might themselves use it deferentially—deferentially to *future* users in *their* community. This is suggested by Putnam's statement that 'the key point is that the relation $same_L$ is a *theoretical relation*: whether something is or is not the same liquid as *this* may take an indeterminate amount of scientific investigation to determine' (1973: 702)—so the proposal is that users of 'water' might use it deferentially to future users because they fix its reference as 'stuff *of the same kind as* these samples' and they use 'of same kind as' deferentially to future users. So the communities on Earth and distant Twin Earth might refer respectively to H_2O and XYZ when they use the term 'water' because the introducers of the term on Earth intended its reference to be that which future users belonging to their community would pick out as 'water'—as did the introducers of the term on Twin Earth.

This is vague, of course. But I need not make it more precise. The point I wish to emphasize is merely that if Kripke's proposal can be suitably understood, elaborated, or modified to secure that the initial baptism does contain an identifying reference to a natural kind, as his account requires, given the parallelism he insists on between the introduction of proper names and that of natural kind terms, then this will merely ensure that at least the initial uses of a natural kind term are in accordance with the Fregean account.

So far I have argued that Kripke's picture of the introduction and transmission of natural kind terms is not inimical to the Fregean account. The Fregean can, similarly, accommodate his thesis that natural kind terms, thought of as proper names of kinds, are rigid designators, by appeal to the convention suggested earlier as the explanation of the rigidity of proper names of individuals. The variation in sense from user to user which makes the convention intelligible may, in the case of natural kind terms, be grounded in a difference between experts and non-experts within the community, but just as in the case of proper names of individuals, it need not be. Even if everyone in the community has access to the—known to be imperfect—set of tests for being water, the difference remains between those who have actually encountered the baptismal sample and those who have not—so deference will take place, and that is all that is required for the convention to have a rationale.

The fact that natural kind terms, construed as proper names of kinds, function as rigid designators, and do so for the reason explained, accounts for the difference we perceive between such terms, construed now as general terms, and general terms for non-natural kinds. Something is a tiger just in case it is of the kind *tiger.* The predicate, like the name, varies in sense from speaker to speaker, with some speakers deferential in their use of it to others and all speakers aware that the tests available to the community may be imperfect, since what it is to be of the kind is to be *of the same kind* as the baptismal samples, whatever the tests indicate. Contrast 'bachelor'. Trivially, something is a bachelor just in case it is of the marital kind *bachelor.* But the rest of the story about 'tiger' does not carry over. We therefore do not have to concern ourselves with extending the notion of rigidity from singular terms to predicates and general terms to account for the contrast Kripke rightly perceives between 'tiger' and 'bachelor'.

Just as the Fregean account is consistent with the ascription of essential properties to individuals, so it is consistent with the ascription of essential properties to natural kinds. We can distinguish between saying that Kripke is made of molecules and saying that Kripke is essentially made of molecules. Equally we can distinguish between saying that gold is an element and saying that gold is essentially an element. But we can express this latter difference otherwise. To say that gold is an element is to say that every (actual) sample of gold is a sample of an element. To say that gold is essentially an element is to say that gold is such that it is necessary that any sample of it is a sample of an element. Every sample of gold is a sample of an element, according to the Kripkean picture, just in case every sample of stuff which is of the same kind as the baptismal samples is a sample of an element, where 'being of the same kind as' denotes that relation which must obtain between a baptismal sample and a sample of stuff if the latter is to be of the kind *gold*, or, in other words, the relation between samples which the initial baptisers' intention fixed as the reference of the relational predicate 'is of the same kind as' that they regarded as denoting the relation that had to obtain between a baptismal sample and a sample of stuff if the latter was to be of the kind they called 'gold'. (The worry discussed earlier about how the chemically ignorant initial baptisers can have the identifying knowledge of the kind gold that Kripke's story requires is, of course, at bottom the worry how their intentions can succeed in fixing the reference of this *relational predicate* as a definite relation between samples—in particular, one satisfying the condition that two samples can stand in it only if both are samples of an element if either is, though they need not be samples of the same isotope—when their tests for sameness of kind can no more discriminate between gold [an element] and fools' gold [a compound] than they can between two isotopes of gold.) And gold is such that it is necessary that any sample of it is a sample of an element just in case any sample of stuff in any possible world that is of the same kind in that world as the baptismal samples are in the actual world is a sample of an element in that world. We can thus express the claim that gold is essentially an element either by using a singular term which designates the kind ('gold' or 'the kind of stuff of which these samples are samples') and a modal operator, or by reference to and quantification over samples and possible worlds and the use of a relational expression denoting a four-term cross-world relation ('*x* is of the same kind in *w* as *y* is in *w'* ').

Despite this difference between the ascription of essential properties to kinds and to individuals, the same account can be given of how we can come to know that an individual or a kind possesses an essential property.

I can come to know, but only *a posteriori*, that Kripke is made up of molecules (that is, I can come to know, but only *a posteriori* the truth of the thought I express by the sentence 'Kripke is made up of molecules'). If I can know, additionally, that *if* Kripke is made up of molecules he is essentially made up of molecules, I can come to know, by *modus ponens*, that he is essentially made up of molecules. Whether my knowledge of the conditional is *a priori* or *a posteriori*, my knowledge that Kripke is essentially made up of molecules, thus gained via *modus ponens*, must be *a posteriori* (see Kripke 1971: 180). More carefully expressed, the reasoning will involve *modus ponens* and universal instantiation. Its premises will be (a) Kripke is made up of molecules, (b) Kripke is a material object (say), and (c) if any material object is made up of molecules it is essentially so. In whatever way (b) and (c) can be known (I have no idea how (c) can be known), since (a) can only be known *a posteriori*, the knowledge that Kripke is essentially made up of molecules, gained in this way, can only be *a* *posteriori*. The indispensability of universal instantiation in this reasoning reinforces the point that what primarily comes to be known *a posteriori* is the possession of an essential property by an object, not a necessary truth (the reasoning would remain valid if 'Kripke' in (a) and (b) were replaced by 'the author of *NN*'—but the knowledge then acquired would not be knowledge of a necessary truth). Consider also the example of knowledge that NP is essentially odd, introduced earlier. The argument here is (a) NP is odd, (b) NP is a number, (c) any odd number is essentially odd. The conclusion that NP is essentially odd can only be known *a posteriori* since it can only be known *a posteriori* that NP is odd.

Similarly, I can come to know, but only *a posteriori*, that gold is an element, i.e. that every sample of gold is a sample of an element (that is, I can come to know, but only *a posteriori*, the truth of the thought I express by the sentence 'every sample of gold is a sample of an element'). If I can know, additionally, that *if* every sample of gold is a sample of an element, then gold is such that it is necessary that every sample of it is a sample of an element, then I can infer, by *modus ponens*, that gold is such that necessarily every sample of it is a sample of an element, i.e. gold is essentially an element. Whether my knowledge of the conditional is *a priori* or *a posteriori*, my knowledge that gold is essentially an element, thus gained via *modus ponens*, must be *a posteriori*. (The additional premises here are (b) gold is a substance and (c) if any substance is an element it is necessarily so [i.e. that if samples x in world w and y in world w' are samples of the same kind of substance, x is a sample of an element in w if y is a sample of an element in w'].)

Of course, in this case, unlike that of knowledge of Kripke's molecular structure, there is the question of how I can know, without investigating all actual samples, and, indeed, the whole world, that gold is in fact an element. Here it seems the Kripkean story requires knowledge of another conditional, namely that if any (actual) sample of gold is a sample of an element, all (actual) samples are samples of an element. And this, it appears, can only be known *a priori* if at all. But then it appears that the same must be true of the conditional licensing the inference from 'gold is an element' to 'gold is essentially an element'. And, indeed, this seems to be Kripke's position. (To be absolutely clear, there is nothing, in general, problematic about the possession by objects, or, at least, by kinds, of essential properties that they can only be known to possess *a posteriori* and so can only be known *a posteriori* to possess essentially—at least, if the notion of analyticity is deemed unproblematic. I can stipulate that 'x is of the same marital kind as y' is to mean that x and y are both married men, or both married women, or both spinsters, or both bachelors, or both widows, or both widowers. Given this stipulation, the marital kind to which my next-door neighbour actually belongs [who happens to be a bachelor] is then such that it is necessarily true of it that if someone belongs to it he is unmarried. But this can only be known *a posteriori*. If I fix the reference of the term 'the marital kind *B*', by specifying that it is to stand for the marital kind to which that man [my neighbour] in fact belongs, it is necessarily true of the marital kind *B* that if someone is a member of it he is unmarried. This is unproblematic because it is trivially analytic, given the stipulation, that if two people are of the same marital kind both are unmarried if either is. Again, if a modern-day chemist asserts 'The element, if any, to which *these* samples belong is necessarily such that any sample of it is a sample of an element with atomic number 79', ostending samples of gold, what he says will be true, because it will be analytic for him that samples are samples of the same element [i.e. not an isotope-limited kind] just in case they have the same atomic number. As a chemically uninformed speaker uses the relational term 'same kind' it is not, however, trivially analytic that if two samples are samples of the same kind of stuff, both are samples of an element if either is. Nevertheless, Kripke thinks that it is knowable *a priori* by philosophical reflection.)

However this may be, given (a) that gold is essentially an element and (b) that it can only be known *a posteriori* that gold is an element, it does not follow, on the Fregean account I have been defending, that there is a necessary *a posteriori* truth: that gold is an element, anymore than it follows from the corresponding premises about Kripke that there is a necessary *a posteriori* truth: that Kripke is composed of molecules. In each case what is knowable only *a posteriori* is that a certain contingent thought is true, and what is necessary, or essential, is the possession of a property by an object, referred to by a constituent of that thought—in the one case an individual, in the other a kind. 'It is necessary that gold is an element' is true, for the same reason that 'It is necessary that Kripke is made up of molecules' is true. But in neither case is any thought said to be a necessary truth. Both in the case of the supposed necessary *a posteriori* truths about individuals and in that of those about natural kinds, the Kripkean notion of a metaphysically necessary *a posteriori* truth is the conflation of the notion of an essential *property* of an object with that of an only *a posteriori* knowable contingently true *thought*. Kripkean metaphysically necessary truth is merely the product of (unintentional) sophistry and illusion; it belongs on Hume's bonfire.

NOTES

1. This can also be expressed as: 'for some α, it is true that [α = the author of *W]* and George IV wondered whether [α was a Scot], if 'it is true that' is understood as forming a context within which expressions have their indirect reference, i.e. as an operator which stands for a function mapping a thought onto a truth-value (Dummett 1981, ch. 9).
2. There is a very similar proposal in Burge (1979a).
3. See Dummett (1974: 527): 'Suppose that the causal theory of reference is correct in that it gives an accurate account of the way in which, in problematic cases, it is generally agreed that the reference of the name is to be

determined . . . Then the causal theory . . . merely gives an account of what senses [names] have. The alternative is to suppose that the causal theory gives a correct account of the conditions for a name to have a particular object as its referent, even though, in critical cases, most speakers would repudiate that means of determining the truth-value of sentences containing the name. This . . . would mean that a certain means of determining the truth-value of a sentence might be the right one although it was not acknowledged as such by any speaker of the language. Such an idea would seem to involve the same fallacy as "They are all out of step but our Willie" '.

4. What of sentences containing propositional attitude verbs within modal contexts? 'George IV might have believed that Scott was Polish' is ambiguous between 'it might have been that for some α, α is an (identity revealing) aspect of Scott and George IV believed that [α was Polish]' (which says of Scott that George IV might have believed that he was Polish), and 'for some β, β is the sense of 'Scott' in my idiolect and it might have been that for some α, α is an aspect of Scott corresponding to β and George IV believed that [α was Polish]'. Hence a possible world in which George believes that Scott is Polish is one in which he at least believes of Scott (our actual Scott) that he is Polish. This is one lesson of Burge's work. Counterfactual Oscar does not believe that he has arthritis in his thigh, because his belief is not a belief about arthritis (the actual ailment), though Actual Oscar, who differs internally not at all, does so, because in the actual community to which he defers in his use of the term, 'arthritis' is used as a name of arthritis (Burge 1979b).
5. The Fregean will interpret the form '*X* is essentially *F*' as equivalent to 'for some α, α is an essence-revealing aspect of *X* and it is necessary that [if something is α it is *F*]'.
6. There are in fact two points here, the second substantially more contentious than the first. The first point: 'gold' could have been introduced in a community to name the kind to which belonged certain samples, picked out ostensively, about which, apart from their location, it was known only that they were shiny, hard, and yellow. The community could certainly acknowledge that there might be other hard, shiny yellow stuff elsewhere, which was a different kind of stuff. The second point: they could also acknowledge that stuff that *they* would not be able to distinguish at all from the samples by any of the tests available to them for sameness of kind might still be of a different kind, because they could acknowledge that their tests for sameness of kind were imperfect and incomplete. Again, it is highly plausible that a primitive community in Bengal might introduce a term, say 'tiger', for the kind of large, dangerous, apparently striped creature that comes in the night and attacks them, whilst recognising that there might well be other kinds of large, dangerous, striped creatures they could not distinguish from those they call 'tigers' (they have never caught a tiger or even got close to one). Additionally, it is plausible that the community could employ a notion of sameness of kind whilst recognizing that *their* knowledge is so limited that pairs of things (samples) they could not at all tell apart might be of different kinds, i.e. might not fall in the extension of *their* relational predicate 'is of the same kind as'. These are the intuitions to which Kripke is appealing.

REFERENCES

Burge, T. (1979a) 'Sinning against Frege', *Philosophical Review*, 88: 398–432.

———. (1979b) 'Individualism and the mental', *Midwest Studies in Philosophy,* 4: 73–121.
Chalmers, D. (unpublished) 'Propositions and attitudes: a Fregean account'.
Dummett, M. (1974) 'Postscript', *Synthese*, 27: 523–34.
———. (1981) *Frege: Philosophy of Language*, 2nd edn. London: Duckworth.
Geach, P. T. (1976) 'Two kinds of intentionality?', *Monist,* 59: 306–20.
Hale, R. V. (2004) 'Putnam's retreat: some reflections on Hilary Putnam's changing views about metaphysical necessity', in P. A. French and H. K. Wettstein (eds), *Midwest Studies in Philosophy*, 28: *The American Philosophers*.
Kaplan, D. (1969) 'Quantifying in', in D. Davidson and J. Hintikka (eds), *Word and Objections: Essays on the Work of W. V. Quine*, Dordrecht, Netherlands: D. Reidel Publishing Company.
Kripke, S. (1971) 'Identity and necessity', in M. K. Munitz (ed.), *Identity and Individuation*, New York, New York University Press. Reprinted in A. W. Moore (ed.), *Meaning and Reference*, Oxford, Oxford University Press, 1993. Page references to reprint.
———. (1972) 'Naming and Necessity', in D. Davidson, and G. Harman (eds), *Semantics of Natural Language*, Dordrecht, Netherlands: D. Reidel, 253–355. Reprinted in (1980) *Naming and Necessity*, Oxford: Blackwell. Page references to original.
———. (1979) 'A puzzle about belief', in A. Margalit (ed.), *Meaning and Use*, Dordrecht and Boston: Reidel.
Lewis, D. (1984) 'Putnam's paradox', *Australasian Journal of Philosophy*, 62: 221–36.
Noonan, H. W. (1979) 'Rigid designation', *Analysis*, 39: 174–82.
Noonan, H. W. (1981) 'Names and belief', *Proceedings of the Aristotelian Society*, 81: 93–108.
Putnam, H. (1973) 'Meaning and reference', *The Journal of Philosophy*, 70: 699–711.
Quine, W. V. O. (1966) 'Quantifiers and propositional attitudes', in *The Ways of Paradox*, New York: Random House.

6 Theoretical Identity Statements, Their Truth, and Their Discovery

Joseph LaPorte

According to a celebrated philosophical tradition that has enjoyed prominence for some decades, so-called 'natural kind terms' that were coined before much science was known, like 'oak', 'water', and 'mammal', refer to kinds with theoretically interesting essences. According to the tradition, scientists learn by empirical investigation what those essences are, and express them with theoretical identity statements, the paradigm of which is 'water = H_2O'. Scientifically informed conclusions about kinds' essences are discoveries, not stipulations. Scientists shed light on the way speakers have *all along* been using the term: 'scientific discoveries of species essence do not constitute a "change of meaning"; the possibility of such discoveries was part of the original enterprise' (Kripke 1980: 138; see also Putnam 1975: 224–5). In section 1, I summarize briefly my response to the foregoing tradition (following LaPorte 2004). In section 2, I defend that response.

1. MY RESPONSE TO THE TRADITION

My response to the foregoing tradition is mixed. I maintain that the usual theoretical identity statements (i) are not *discovered* to be true, but, at least sometimes, (ii), they *are* true. I begin with a very brief discussion of the second point, illustrating with biological kinds: these are especially controversial but they provide the best case.

Specialists in the philosophy of biology have often chafed at essentialism with respect to biological kinds, with some justification. Essentialists have not been well informed about biology: accordingly, *some* commonly held essentialist theses should be abandoned in light of contemporary biology. Nevertheless, some forms of essentialism about biological taxa are highly plausible in view of contemporary biological systematics and cladism, the leading school, in particular. Thus, the kind *Mammalia*, as some scientists understand it, has a relatively easily specified essence. It is the kind including, with respect to all possible worlds, *M* plus all and only *M*'s descendants, where '*M*' is a *name* I assign to the nearest common ancestor of horses and echidnas: Mammalia is the *clade* with stem *M*.[1] *Chordata* is the clade with stem *C*, so Chordata includes, with respect to all possible worlds, *C* plus all and only *C*'s descendants. *C* is, in fact, the nearest common ancestor of sea otters and sea squirts. And so it goes.

What about the further claim that biologists' conclusions represent discoveries about the essences of kinds discussed in a tradition antedating modern science? Here I would part with the familiar line. According to that line, scientists have *not altered* the use of key terms in essence-exposing theoretical identity sentences like 'Chordata = the *clade* with stem *C*' to *make* such sentences true; according to that line, scientists have just *found* such essence-exposing theoretical identity statements to have been true *all along*. Contrary to this received position, I hold that scientists change the meanings of kind terms by *stipulatively deciding* how to use traditional terms, in response to conceptual disruption in the relevant inquiry, when science develops. The problems for claiming that we discover essence-exposing theoretical identity statements about biological kinds specifically is illuminated by a look at three phenomena or alleged phenomena. They are (1) supposed corrections of past speakers, (2) conflicting scientific characterizations about how to delimit taxa, and (3) the need for stipulation to resolve the breadth of kinds even given one agreed-upon scientific characterization. Closely related points apply to the other widely discussed group of natural kinds, chemical kinds, which I set aside for the moment.

With respect to the first issue (1), Kripke says that empirical research has resulted in scientists' 'discovering that "whales are mammals, not fish" is a necessary truth' (Kripke 1980: 138). I think Kripke is wrong here. We *might* have said, in view of all of the relevant empirical information, 'whales *are* fish'. We might have said this on the grounds that *fish* turns out to be an *unnatural* category, from the perspective of biological inquiry, by virtue of including sharks and whales to the exclusion of cows. Had we drawn this conclusion, we would have *dropped* the use of the term 'fish' from scientific discourse instead of *keeping* the term for a natural group by excluding whales from the term's extension.[2]

There are other grounds on which we might have said 'whales *are* fish'. We might have said, 'even though *fish* is a *natural* category, it turns out to include whales *as well as* cows and other tetrapods like ourselves'. You might balk at the idea of extending 'fish' to cows but there is precedent for such dramatic extension: surprisingly enough, scientists now generally say 'birds *are dinosaurs*', though it took a while to catch on. Birds turn out to be related to stereotypical dinosaurs of an earlier age in the way that cows and other tetrapods are related to stereotypical fish of an earlier age: they are descended from them. We still say 'whales are not fish', as we did before we knew about the relevant evolutionary relationships; but we do not still say 'birds are not dinosaurs', though our information about relationships has changed in relevantly similar ways.[3] Both ways of handling the relevant change in scientific information seem acceptable: in general, there are different possible responses to conceptual disruption (see note

3) and no *one* is required by the use of prescientific speakers, and hence discovered to be correct.

The failure of discovery becomes clearer when we raise the second sort of consideration that I have mentioned: (2) conflicting scientific characterizations about how to delimit taxa. Sometimes issues raised by (1), correction, do not arise: sometimes everyone agrees from the start on what actual organisms belong to the kind. But even then, the discovery picture is mistaken. Scientists disagree about what *makes* an organism belong to its kind: whether interbreeding determines the matter, say, or a finer criterion, whether common descent is all important or whether absence of evolutionary change is required. If and when one camp wins out,[4] it is hard to believe that it is discovered to win out, since each camp seems to break organisms into natural groups that do justice to earlier speakers' naming (LaPorte 2004: 70–83).

Finally, we reach (3). Even without the foregoing scientific disagreement about the *nature* of essences, the *scope* of an essence's *extension*, so to speak, is settled by stipulative measures. So if the *bears* are to be a stem population and its descendants, instead of a group with some other sort of essential nature, then scientists will pick some stem population or group as the key to the kind's essence; but they might have picked another stem, which would have done as well. The stem chosen for bears might be some species far back in the lineage leading to extant bears, or it might be a relatively recent ancestor, depending on the systematist's preferences. Which stem is selected as the key to the *bears*' essence determines whether some species are or are not *bears*: hence, the giant panda, which separated from the bear lineage early on, could go either way (LaPorte 2004: 84–5).

I have discussed theoretical identity statements exposing the essence of biological kinds. The lessons that I have drawn apply more broadly. We are supposed to have discovered the truth of the theoretical identity statement 'water = H_2O'. Putnam (1975) famously says that if we ever visited a distant planet Twin Earth and found a water-like substance with a long chemical formula XYZ, we would say that the new stuff is not water. Kripke (1980: 128–9) says that if we found H_2O without watery properties somewhere, we *would* count it as water. Such intuitions are thought to indicate the discovered truth of the theoretical identity statement 'water = H_2O'.

The history of the term 'jade' relevantly resembles Putnam's story of Twin Earth. But if my information is right, *real* speakers did not evince intuitions of the sort that Putnam is committed to saying that they would. In the case of 'jade', speakers encountered a new substance with properties similar to those of a widely known and valued substance that they had been calling 'jade'. The new substance, jadeite, had a completely different microstructure to the old substance, nephrite. Even so, speakers responded by applying the word 'jade' to the new substance. So today, jadeite counts as belonging to the extension of 'jade'.[5]

A term could be extended to newly discovered stuff with the right *superficial properties* but the wrong microstructure. So, too, a term could *fail* to be extended to stuff with the right *microstructure* but the wrong superficial properties. When scientists found that blue sapphires are comprised of the same mineral that other stones like red rubies are comprised of, namely the mineral corundum, they did not then apply 'sapphire' to rubies and other stones comprised of this mineral. Instead, they interpreted the term 'sapphire' as a colour-restricted varietal term. Nor does their decision seem to have been *discovered* to be right; we *could* as well have extended 'sapphire' to the whole mineral. We did that with 'topaz' when we found that the underlying mineral is not all the originally observed colour, yellow. We'd been calling other colours of topaz by other names but we now call them 'topaz'.

I have been looking at the role of superficial properties in delimiting a kind in different possible worlds: that is, their role in essences. Even if we restrict ourselves to consideration of *microstructural* properties' role in essences, it seems to me that our theoretical identity statements could often as well have been passed up in favour of other theoretical identity statements that match our vernacular kinds to different microstructural essences. Thus, it seems to me that *some* H_2O that is now counted as water, namely D_2O, could have been counted nonwater without error. D_2O is a form of H_2O with certain salient properties that distinguish it from ordinary H_2O: drinking it will not sustain our biological processes, for example. D_2O behaves differently than ordinary H_2O in nuclear reactions, its ability to absorb radiation (to which Nigel Leary calls my attention: p.c.), and across a range of other matters (see LaPorte 2004: 106–7).

The view that scientists' conclusions about the nature of our kinds are discoveries is wrong: that is my basic claim. There are consequences of philosophical importance.[6] Rather than explore the details of such consequences, though, I will defend the basic claim itself, in view of salient objections lodged against it.

2. WORRIES

Worries over my preferred view can be outlined naturally by topic. I argued in section 1 that theoretical identity statements (i) are not *discovered* to be true, but that, at least sometimes, (ii) they *are* true. The worries to be addressed correspond. With respect to (i), discovery, there is the objection that theoretical identity statements *are* discovered to be true after all, and further that even if they are *not* discovered to be true, that does not matter much. With respect to (ii), truth, there is the objection, from the other direction, that theoretical identity statements are not even *true*, let alone discovered to be true, and further that even if they *are* true, that does not matter much. The foregoing outline permits me to address two important interlocutors by turn: Alexander Bird presses the objections to be taken up in section 2.1 and John Dupré presses the objections to be taken up in section 2.2. I begin with (i), discovery.

2.1 Whether Theoretical Identity Statements are *Discovered* to be True, and Whether it Matters

Bird (2007) doubts my position on both issues.

Whether Theoretical Identity Statements are Discovered to be True

Bird has a higher regard than I do for the relevant discovery account concerning theoretical identity statements from the vernacular. Let me consider his arguments from chemistry and biology in order.

First, Bird takes up terms like 'jade', 'ruby', and 'topaz'. For him, it might be appropriate to count it a *discovery* that 'ruby' and 'sapphire' designate just one colour, or a limited range of colours, of stones of the composing mineral, and that 'topaz', by contrast, designates *all* colours of its mineralogical composition. I regard these conclusions as the result of historical accidents, and not discovery. 'Ruby', 'sapphire', and 'topaz' were all originally applied to stones of just one variety, not the whole mineral, but whether the term should be restricted to that variety was unclear in the centuries before the discovery of mineralogical structures.

Bird points out that the term 'corundum', which designates the *mineral* of which ruby and sapphire are *varieties*, was *already available* for the mineral *species*, so it is only natural that it should go on designating that *species*, while 'ruby' and 'sapphire' should go on designating a *variety* of the species. Here we seem to have an empirical disagreement. I agree that the name 'corundum' was around for C. F. Greville to choose when he unlocked the nature of the mineral. But 'corundum' was *not* already a mineral term. It was itself applied just to a *variety* of the relevant mineral, and speakers might as easily have decided to keep it that way. So again, it would appear to be an accident that this term, rather than some other that had *also* been applied just to a variety of the mineral, was extended to designate the whole mineral.

'Corundum' was a foreign term passed on to Greville as a label for some obscure samples of impure, industrial-grade corundum that he received for purposes of analysis (Hughes 1990: 1): the stuff was allegedly used by natives of India for purposes of polishing hard things, including 'rubies' (Greville, de Bournon, and Oakley 1798: 404). Greville found the connection between the samples of industrial-grade matter passed on to him as 'corundum' and the lofty rubies and sapphires that turned out to be the same mineral. It isn't clear that a term for industrial-grade specimens should go to the mineral in general, rather than to industrial-grade corundum. Further, had Greville not received his samples of industrial-grade 'corundum', which was an obscure material that might more easily than not never have come his way (Greville, de Bournon, and Oakley 1798: 408), he might still have discovered the affinity between rubies and sapphires, in which case it is far from clear that he would or should have kept 'ruby' and 'sapphire' for varieties. He might well have used 'sapphire' for the mineral: 'sapphire' was applied to colours other than blue by contemporaries of Greville. Indeed, there was already a tendency on the part of some authorities to apply the label 'red sapphire' to rubies (Greville, de Bournon, and Oakley 1798: 419), although the tendency is no longer to be witnessed today: speakers today apply 'sapphire' to all colours of the gem except red. The lesson of the foregoing historical details, I suggest, is that historical accidents, not discovery, are responsible for speakers' having decided to use 'corundum' as an inclusive term for all of the mineral whose structure Greville discovered, and to restrict 'ruby' and 'sapphire' to their present-day varieties.

Bird has further worries. Since terms like 'sapphire' and 'jade' are names of gemstones, he says, 'there are interests competing here with scientific ones' (2007: §3).[7] I accept this observation. I would extend it to other kinds from the prescientific vernacular too. People name kinds that are salient to them and their interests: cf. 'water' or 'fish'.

To Bird, the involvement of lay interests, with respect to terms like 'jade', suggests that ordinary speakers might be expected to draw the wrong conclusions about essences: hence, 'jade really was just nephrite but the makers of jade artifacts decided it was in their interests' (2007: §3) to alter the meaning of the term. I would say, on the contrary, that such tension caused by competing interests issues in a species of semantic indeterminacy or vagueness: namely, open texture, which I understand to be vagueness yet to be exposed by discoveries that bring to light borderline or unclear cases for a word's application (LaPorte 2004: 188, note 4; see the related discussion in this paper at note 10). There is an inclination to judge the proper use of a term in *each* of two or more *competing* ways, at least one of which represents scientific interests and one of which represents cultural interests: but no one of these ways of resolving the proper use of the term trumps the others unless and until a decision is made to refine the use of the term in one or another direction. That removes the tension.

I cannot rule out definitively, with any investigation of speakers' behaviour or judgments like this, the possibility that there was a straightforward use instead of vagueness for 'jade': similar words apply to other terms in natural language, like 'tall' or 'hot'. But the intuition that there was a straightforward use of 'jade' for the original mineral baptized seems to be motivated by the very theory that the intuition is supposed to support. Claims about whether this or that sample, e.g. of XYZ, could be such and such, e.g. water, are generally supported by intuitions about what speakers *would say* in Twin Earth–like scenarios like the ones that I have been talking about. And what speakers would say, indeed did say, in the case of jade turned out not to support the theory.

'Water', like 'jade', was vague before modern science, or so I claim. Speakers might have concluded that 'water = H_2O' is false, and been no more wrong than they were in concluding that 'water = H_2O' is true: that is because speakers could have come to count D_2O as nonwater. Here the vagueness arises even if it is clear, unlike for 'jade', that the relevant essence is microstructural.

Several able philosophers have suggested to me, in and out of print, that 'from the point of view of chemistry' it would have been a mistake to say that D_2O is not water. D_2O reacts chemically with other substances just as ordinary H_2O does: it is a solvent, for example, and the elemental ingredients can be brought together to form the compound and then decomposed by chemical means—D_2O is, after all, a form of H_2O. It would be highly unnatural, I have been told, for earlier speakers to have concluded, after discovering the existence of D_2O, that this is not *water*.

I do not agree: it is true that if the only relevant point of view concerning how to use the term 'water' properly were the view of *chemistry*, it would have been a mistake. But with respect to a term like 'water', *chemical* properties are not decisive and other properties, including biological properties and physical properties, matter too. Visitors to Deuterium Earth, where the water-like stuff is D_2O (LaPorte 2004: Ch. 4; cf. Putnam 1975) might find that what they are tempted to call 'dwater', D_2O, has its own structure and strange properties. It is not even potable. It is not dangerous to be around for what it emits, as radioactive substances are (Bird makes this point), but *it does not perform the life-supporting functions of their 'water'* inside the body. That presents reasonable grounds for earlier speakers, in the wake of D_2O's discovery, to have concluded not, as we happen to have done, 'D_2O is water', but rather 'D_2O is not water, though it *is* one form of H_2O'.[8] We cannot say that this alternative conclusion would have been wrong and that the conclusion that we ended up with was right.

Related remarks apply to the chemical elements. As Bird observes (2007), the ancients had pure samples of a handful of elements, including gold, which has been used since Neolithic times in jewellery. They also had carbon, which could be found in charcoal. But were the ancients' names for substances really designators of our *chemical elements*, unbeknownst to them? It seems unlikely. Speakers found that charcoal is made of the very same chemical *element* as diamond, though it has a different molecular structure. Still, speakers did not say 'diamond turns out to be charcoal' or vice versa, though they might have done so without straightforward error, it seems to me, since they might have interpreted 'charcoal', say, as a general name for the element when they discovered the affinity between diamond and charcoal. As it happens, in view of differences between charcoal and diamond, speakers have kept distinct rubrics for them, declining the opportunity to use 'charcoal' as a general name for the element. Speakers who first learned of the elemental status of gold samples might with equal justice have said that any matter comprised of the element of atomic number 79 that is not at all like familiar gold in its properties, having taken on a new structure, would *not* be in the extension of the ancient term 'gold', which is accordingly not a general name for the element.

Let us turn to Bird's claims about biology. He proposes an account of how biological terms should be reassigned to extensions in the face of new empirical results. That account is supposed to underwrite discovery by showing that there is no reason to stipulate new uses for older terms as conceptual disruption occurs: the older use has clear application, contrary to what I have said, and scientists' reports honour this older use. Bird calls this proposed account 'TAX'. In essence, TAX requires that after conceptual disruption, speakers keep the extension natural, if possible, by leaving in a majority of clearly 'typical' members and preventing a majority of clear 'foils'. If the extension cannot be kept natural in this way, speakers are to strip the term of official recognition and let it designate an artificial kind.[9] On its face, TAX is highly plausible. But it does not, it seems to me, represent 'requirements of the way taxonomy works' or upset the idea that any case of correction that I discuss 'could have gone another way' (Bird 2007: §3). Bird challenges me to say more about why. I will try.

As I see it, TAX is generally unable to put a stop to vagueness: what counts as a 'typical' member or foil is unclear: that generates vagueness. For instance, it is *not* clear that the word 'reptile' should fail to come to apply to birds, which are foils, in the way that Bird suggests: many systematists have refrained from applying 'reptile' to birds (see, e.g., Hull 1998: 274) but many others have counted birds as *reptiles*, and in this way preserved the term for a natural group (see, e.g., 'Those Diverse Diapsids' 1994–2009). Similar examples to illustrate the same point are close at hand (see note 3).

The foregoing paragraph raises indirectly a further problem for TAX: it would invite discord. When scientists defer to something like TAX, different scientists reconstruct in different ways in response to disrupting empirical information, as they have with 'reptile', some having dropped the term and others having extended it to cover birds, in view of the surprising news that birds are so closely related to traditional reptiles. As a result, different scientists might easily end up speaking at cross purposes. Accordingly, different groups of biologists have proposed well-known measures that would do more than TAX would do to ensure that there is a *definite answer* as to how a term is to be used in the event of unforeseen disruption. Different sets of such measures are incorporated into different codes. I will discuss one such code (Cantino and de Queiroz 2007) that is now followed by many systematists, a code that does not seem to be just wrong or unworkable in view of the biological facts: Phylocode. This code would cross sharply with the requirements of TAX over various cases. For example, the code would allow for 'Rodentia' to be defined as the clade with stem *R*, the closest common ancestor of guinea pigs and mice (for discussion of a similar proposal for 'Rodentia', see Wyss and Meng 1996: 560–1): this definition would allow and even force scientists to accept guinea pigs as rodents, even if guinea pigs turn out to be more closely related to horses and us than mice, by extending 'rodent' to horses and us (Bird appeals to TAX to reject that conclusion). My point here is not that promoters of Phylocode have discovered that the theoretical identity statements that Phylocode honours have been discovered to be true; consider only that the code allows for discretion in choosing stem groups like the foregoing. Nor is my point that systematists who reject TAX, including promoters of Phylocode, have discovered that TAX and its verdicts with respect to classification are wrong. My point is just that the many biologists who reject TAX and its verdicts are themselves not wrong in doing so.

I would reject TAX as a proposal to explain how terms are after all handled in a way that frustrates claims to open texture, then. TAX would not preclude the sort of vagueness I talk about, if it were implemented. And it is not generally implemented; nor do scientists seem wrong for deviating from it. In general, it seems hard to avoid the conclusion that scientists refine away open texture at the prompting of conceptual disruption, and that this process undermines claims to discovery.

I would like to emphasize, as I prepare to draw a close to my discussion of open texture and refinement, that I have never proposed that, since vagueness in earlier terms can be refined out in different ways, decisions about how to use terms in view of disruption is all willy-nilly. There *are* serious considerations on behalf of different possible refinements. So it is natural that speakers should often stand by their guns in arguments that turn on the proper use of vague terms when, despite speakers' firm convictions that they are conforming to ordinary standards, the speakers would really have to refine in order to have a claim that overcomes objections from gainsayers. 'By golly, that drink is *hot*—I can sip it only slowly!': such considerations carry force. But I think that this force can be acknowledged along with the vagueness that would prevent this force from trumping considerations on behalf of competing refinements: 'the drink is only warm, since I turned off the burner some time ago'.

There are different ways to accommodate both the *force* of considerations on behalf of what amounts to a refinement and the observation that these considerations do *not*, despite their force, overcome competing refinements and thereby preclude vagueness. One way to accommodate is to say that no disputant is quite right or quite wrong, in her affirmation of what would amount to a refinement. I can produce good reasons to support my claim 'the drink is hot', but you can produce good reasons to support your claim too: so it might then be that our competing reasons cancel each other out. Or perhaps the force of our reasons is overwhelmed by other complex considerations about language use. In either case, the result might be that we are both somewhat right and somewhat wrong.

Another way to accommodate the force of our reasoning while still recognizing the vagueness that prevents my reasoning from trumping yours, or vice versa, would be to say not that *neither* of us is quite right, but that we are *both* right. Here we could appeal to something like David Lewis's 'rule of accommodation: what you say makes itself true, if at all possible', barring unusual trouble, 'by creating a context that selects the relevant features' as decisive in determining the truth of your claim, 'so as to make it true' (1986: 251). The relevant features would be those that you adduce on behalf of your claim. Accordingly, you are right in supposing, as you take the drink off a cool burner, 'it is no longer hot'; but I am also right in supposing, as I sip the drink, 'it is hot'. Our different standards apply in different contexts.

The application to kind terms is straightforward. After the relevant empirical advances raise the question of whether D_2O should count as *water*, different speakers with different convictions could be right in the context they set, with regard to only-apparently incompatible answers, before a decision to go one way or another. Someone might insist, 'D_2O is best called "dwater": it is not *water*, the substance that sustains us—D_2O *wouldn't* sustain us'. In this way the speaker could set a standard in which D_2O does not count as what she calls 'water', but only what she calls 'dwater'. Alternatively, a speaker could emphasize chemical affinities between D_2O and H_2O to set another context and thereby allow for D_2O to count as what she is calling 'water'. Acceptable contexts change over time for terms that, like 'water', were earlier marked by open texture, as I have emphasized: after a time we can inform others with confidence, in just about any context, 'D_2O is water'.[10] At that point, we might appeal, if questioned, to standard reference sources and the authority of common use: indeed, we could now appeal to just such support for 'D_2O is water' (see LaPorte 2004: 189, n. 8). Those who would benefit from the division of linguistic labour and avoid the eccentric use of words conform their speech accordingly.

Whether it Matters that Theoretical Identity Statements are Discovered to be True

I maintain that we did not discover that 'water = H_2O' is true: the sentence was *not* straightforwardly true as it was used by earlier speakers. The meaning has changed. Whether or not I am right about *this* point, Bird would have a remaining objection. The objection is that even if theoretical identity statements are *not* discovered to be true, that does not *matter* much because certain *non*-identity statements *are* discovered to be true and *they* do the same job of informing us about natural, interesting essences that theoretical identity statements do: theoretical identity statements' work for essentialism is therefore superfluous.

Bird raises this issue in the context of discussing the species problem: he does not dispute that there is a multitude of species concepts, no one of which can be discovered to be correct. So he concedes 'that antecedently to the choice of a specific concept, we cannot be in a position to have knowledge of essences' (Bird 2007: §4). But, he goes on, even if theoretical identity statements elude us, so that we do not have a characterization of the *essence* of a species (which would require an account of metaphysically necessary and sufficient conditions for belonging to the kind or being the kind), we may affirm a great many claims about *essential properties* (these would be metaphysically *necessary* conditions for belonging to the kind or being the kind): 'One can be ignorant of what exactly constitutes essence while nonetheless knowing certain essential facts' (Bird 2007: §4). So, even if one could not specify the essence of the Ceylon spiny mouse, one might know that the Ceylon spiny mouse is not identical to the African pygmy mouse, or that neither of the just mentioned species belongs to the rabbit kind or clade: facts like the foregoing stand up, notes Bird, 'on any plausible species concept' (Bird 2007: §4). Hence, such 'facts about essence can be discovered' (Bird 2007: §4). We could add similar examples from chemistry: we discovered that water is not a form of coal, say, and that this is necessarily true.

What can be said in response to this worry? I would concede a fair amount to Bird: 'genuine discoveries of essential facts' (Bird 2007: §7) are possible and indeed commonplace. Still, I would reject Bird's suggestion that our not having discovered the truth of now-accepted theoretical identity statements is somehow not interesting or not 'problematic' (Bird 2007: §4). As Kripke indicates, scientists work hard to get right theoretical identity statements, which are especially interesting claims: accordingly, our affirmation of them marks a salient case of conceptual development. It is therefore interesting if, in the course of this development, *meaning* change attends theory change, in the way that I argue, and that as a result of meaning change, these interesting affirmations about identity do not really amount to reported *discoveries* that the relevant statements are true. I will say more next.[11]

2.2 Whether Theoretical Identity Statements *are* True and Whether it Matters

We do not discover that essence-exposing theoretical identity statements are true. So I say anyway. But *are* they true? At least one philosopher with a formidable familiarity with the relevant biological discourse, John Dupré, doubts that theoretical identity statements of the sort that I have cited, like 'Chordata = the clade stemming from stem *C*', where *C* is a population, really are true and indeed necessarily true, on their current meanings, as I claim. Further, Dupré wonders what good such theoretical identity statements could be for *essentialists* even if they *are* true: so even if such statements are true, that does not matter. I will here address these claims in order.

Whether Theoretical Identity Statements are True

Is 'Chordata = the clade stemming from stem C' even necessarily true? I say that it is, for a systematist who uses 'Chordata' as a designator for the relevant clade, since that clade is to be identified by its stem, *C*, from possible world to possible world: anything descended from *C* in any possible world *w* is in the clade in *w* and anything not descended from *C* in *w* is not in the clade in *w*. Of course, even if the reader *agrees* that the statement is true for any systematist who *does* use 'Chordata' in the right way, she might worry that no systematist *does*, so that the relevant identity-expressing sentence is not true for any speaker. Dupré is doubtful that any systematist uses 'Chordata' in the right way. In order to use 'Chordata' in the right way, scientists would have to use that term, when considering other possible worlds, for what has the right stem. But to catch scientists openly considering counterfactual worlds is not easy: 'I have never heard any biologist discussing anything of the kind', Dupré (2004) testifies.

Scientists do, however, discuss these matters openly. The following is taken from Mark Ridley, a well-known evolutionary biologist.

> If a chimpanzee population were by some freak of nature (and this is a thought experiment, to make a point, not a realistic hypothesis) to produce offspring . . . possessing six legs and a jointed exoskeleton, the phylogenetic branching pattern would not be changed at all. (Ridley 1989: 2)

That is, chordates would still have *their* stem, and insects theirs. Accordingly, Ridley goes on,

> The chimpanzees would not suddenly have been catapulted into the insects. They would phylogenetically stay where they were, and only certain taxonomists' ideas about how to *recognize* a chordate would have altered. (Ridley 1989: 2; emphasis added)

Claims about 'counter-to-fact' worlds (de Queiroz 1995: 225) and, in effect, essences are entertained especially clearly when scientists address thorny issues concerning the proper way to define taxonomic terms (see, e.g., the exchange between Ghiselin 1995 and de Queiroz 1994, 1995). The indication is that theoretical identity statements like 'Chordata = the clade stemming from stem C' are necessarily true, as at least some scientists use those sentences.

Whether it Matters that Theoretical Identity Statements are True

Even if I am right that supposedly essence-exposing theoretical identity statements like 'Chordata = the clade stemming from C' are necessarily true, on their current meanings, Dupré (2004) wonders what *good* such theoretical identity statements could be for *essentialists*. Dupré would object that the so-called 'essence-exposing' theoretical identity statements support at best a worthlessly weak version of essentialism. I do not think that the essentialism that I defend is weak in this way. Granted, it is not as strong as the sort of essentialism that Dupré has addressed critically in his published research (see, e.g., Dupré 1986) but it does have substantial metaphysical bite.

Let me address two interesting worries along these lines. One worry is that my essentialism concerns only *kinds* or *taxa*; it is not strong enough also to concern *individuals*. The species *Panthera tigris* has an essence. But that is not to say that individual tigers are essentially members of *Panthera tigris*. It is only to say, roughly, that there is something that it *is* for an entity *to be* or to be a *member of Panthera tigris*. There is an essence E such that in any possible world w, x is a member of *Panthera tigris* in w if and only if x has E in w. Whether x exists without E in possible worlds in which x is *not* a member of *Panthera tigris* is a different matter. Dupré remarks, 'Biological essences that have nothing to say about individual organisms strike me as toothless' (2004). As I've said, I find more bite than that.

Perhaps I can convey the relevance of kind essences better by switching examples, in favour of chemical elements. There is a kind whose essence is to be the element with atomic number 82, i.e. the element with 82 protons. There is another kind whose essence is to be the element with atomic number 79, i.e. the element with 79 protons. Suppose that we agree to use 'lead' and 'gold' for these respective kinds. If you think this involves some refinement of English, as I do, so be it.[12]

It seems clear that the essence of a chemical kind like this is interesting: it matters that gold could fail to have its yellow colour, say, or its use in jewellery, but not its status as a chemical element. But again, individual essences are a separate matter. Take an atom of lead: with a particle accelerator we could knock out three of its 82 protons, leaving it with 79. The remaining atom would be gold, not lead. But whether that remaining atom would be the very same one, having undergone a loss of protons, or whether it would be a different atom entirely, remains to be addressed. Perhaps it would be a mistake (as Enç [1986] would insist) to say that our individual atom of lead would go out of existence if it lost protons. Even so, it remains correct to say that the essence of the *kind lead* is to be the element with 82 protons and that the essence of the *kind gold* is to be the element with 79 protons: it remains correct to say that gold could fail to have its yellow colour or its use in jewellery but not its status as a chemical element. Such observations concerning the essences of kinds do not lose their truth or interest, even if we cannot make any related observation about any *member*'s essential possession of the kind's essence.

Let me take stock. Dupré is sceptical about whether theoretical identity statements reveal anything interesting about essences, and I have addressed one interesting source of his doubts: that such essences would principally concern kinds or taxa and not members, like organisms. But that source of doubt that I have addressed is not Dupré's only important source of doubt. I claim that scientists *refine our use of theoretical terms*, in affirming theoretical identity statements, thereby making the relevant sentences true:[13] this naturally gives rise to another type of worry about my essentialism. Dupré remarks that 'the possibility that scientists may *construct* necessary truths by defining their terms . . . says little of interest about essentialism' (Dupré 2004; my emphasis).

The precise nature of Dupré's worry is not stated but plausible candidates suggest themselves. The worry might be that on my view *essences* are *constructed* or *invented* and hence that the spirit of essentialism is lost. Or the worry might be that since, on my view, we stipulate essences for terms, the result is that a term is somehow matched to an artifact, so an *artificial* rather than natural kind. Let me address these worries briefly.

First, I confront the worry that on my view essences are constructed or invented. I wish that I had confronted this position more directly (in 2004) in order to reject it. I am typically summarized in a way that suggests this position but I do not embrace it. And I would reject it (as I indicate in the book's second note: 2004: 175). Usually the mistaken characterization is harmless, given the characterizer's path of inquiry, but the matter should be clarified, where worries concern the seriousness of my essentialism. My position is not that we *construct* or *stipulate* or invent essences. It is rather that we stipulate that our names *attach* to certain essences, rather than discover that our names have attached to those essences all along. In my view, no human effort could make whales fish; but human effort *could* assign the *word* 'whale' a use on which it is correct to say, 'whales are fish'. No human effort could have made D_2O fail to be water; but human effort could have assigned the *word* 'water' to a more restricted extension than it has, so that D_2O, which does in fact belong, would *not* have belonged. Had that happened, the sentence 'water = H_2O' would have been false, since extension of 'water', along with the corresponding essence, would have been more restricted and excluded some types of H_2O. But if the statement *had* been false in this way, then the sentence would have addressed a *different* kind which we would have *called* 'water' and *not* the *same* kind with an *alternative essence* (see Devitt 2009: 54 for a related discussion to similar effect).

In general, on my view, there are *truths* about *essences* of chemicals, clades, and the like, i.e. about what it is to be a given chemical or clade, and these truths are not stipulated or constructed or invented. But when scientists come to affirm theoretical identity statements like 'water = H_2O', they refine the use of a word like 'water' to assign that word to the right essence and thereby *make the sentence true*; another reasonable assignment, given open texture in the prior use of the term, might have made the sentence *false*, by causing the word to come to be matched with a different essence. Then the sentence would have come to say something about a different kind altogether, which has its own essence; *that* is why the *subject of the sentence* would have come to have had a different essence. The sentence would not have come to say anything about the kind that *we* actually refer to with 'water'; we could hardly have come to say of *that* kind that it has a different essence and been right, no matter how we had chosen to assign our words to kinds.[14] Perhaps Dupré misunderstood me on this point and that is why he thinks that my essentialism is thin and uninteresting. At any rate, if one were to misunderstand, one would naturally be led to the conclusion that my essentialism is thin.

There is also the question of whether, given my affirmation of stipulation, we assign *natural* extensions to our kind terms. If we *stipulate*, it might seem that the result is artificial in some way: perhaps we assign unnatural essences to our kind terms. So goes a prima facie worry. I would confront this worry by denying that stipulation need be associated with artificiality. Suppose that you create or discover a brand new mineral, never before recognized. You *stipulate* that it is to be called 'Mineral M'. Here you *stipulate* the use of 'Mineral M' rather than discover how that term was used all along. But you still assign the term to a *natural* referent rather than an artificial one.

In response to Dupré, then, I maintain that the essentialism that I embrace is worthy of the name. It is not as potent as some forms of essentialism that Dupré has targeted in well-known work and, to my mind, discredited (e.g. in Dupré 1986); but it is still plenty substantial.

This is a natural place to revisit some qualms I broach at the end of section 2.1, to achieve a more thorough resolution. Bird, like Dupré, suggests that the truth of theoretical identity statements is not important for essentialism. Bird's reservations are not Dupré's; for Bird, essentialist truths underwritten by theoretical identity statements are substantial enough, but they are *superfluous*. That is because even if we do not know the truth of any theoretical *identity* statement concerning some species of mouse, say, we might know some truths about the mouse species' essence: that this species is not and could not have been a species of *rabbit*, say. In view of such information about essence, Bird seems to maintain that our ignorance about theoretical identity statements, where it exists, is not much of a loss, or 'especially problematic'. I do not agree.

There is a reason scientists try to get clear about theoretical identities. A theoretical identity statement tells us what it *is* to *be* a mouse of this or that species or a rabbit. Such information is clearly richer than the information that this or that species could not have been such and such other species. The information that this or that species could not have been such and such other species tells us little about what it takes to *be* the relevant species: i.e. what the *essence* of the relevant species is. The information conveyed in a theoretical identity statement, by contrast, is precisely what the relevant essence is.

Not surprisingly, scientists *care* about what a taxon fundamentally is (hence, the memorable title of an article published many decades ago, 'What, If Anything, Is a Rabbit?': Wood 1957). In effect, they care about what a taxon's essence is, though they avoid the *word* 'essence': so they care about theoretical identity statements about the taxon, which convey the relevant information. One reason that scientists care about the information at issue is that without it, theorists can talk past each other: paleontological study, say, can be caught up in 'endless and fruitless debate on the question of whether or not certain fossils should be considered "reptiles", "birds", "mammals", or "men", and when those groups evolved' (Hennig 1965: 114). Here is one good reason to put forward identity statements specifying the essences of the foregoing groups.

3. CONCLUSION

Despite important and natural worries broached by capable philosophers, the claims of section 1 seem solid. Essentialism remains plausible in view of current science and current biology in particular. However, the usual views according to which scientists *discover* and *report* the truth of theoretical identity statements *conveying* the essence of kinds that speakers have all along been discussing, with the relevant subject terms, are mistaken.[15]

* * *

I thank Helen Beebee and Nigel Leary for comments leading to improvements, as well as participants at the Nature and its Classification conference in Birmingham, UK: especially Alexander Bird and John Dupré, my commentators there.

NOTES

1. See LaPorte 2004: 11–12. Different workers differ slightly in their use of 'clade': I include the stem, but little rests on it here. More bothersome are different uses of a word like 'Mammalia' or 'fish': sometimes when I say that an organism does or does not belong, I am talking about what certain scientists would call by that name, other times what some group of earlier speakers would call by that name, and so on. I claim that such words' use changes over time, so what counts as belonging to the extension will vary depending on the use in question. In general, to keep the presentation readable, I cue the reader with italics where caution is needed and let context clarify the use in question.
2. There is precedent for this course of accommodating a deviant group by giving up on the scientific status of the taxon: see LaPorte 2004: 69. A complication here is that according to cladists, whose taxa reflect just genealogy, the *un*accommodating reaction of *ex*cluding whales from the 'fish' category would not, contrary to what Kripke suggests, preserve the good scientific name of the fish anyway, at least officially; but the present example of correction presupposes that experts recognize the fish (without whales) as a natural biological group, and that they inform lay speakers accordingly. Expert recognition might be found with specialists who acknowledge the existence and usefulness of such a group, for *some* biological purposes, without giving it recognition in any formal general reference system (Nelson suggests such

a position: 2006: 1–2). Or recognition might be found in a venerable non-cladistic tradition, which accords formal recognition to 'Fishes' or 'Pisces': some contemporary systematists, those who embrace so-called traditional 'evolutionary taxonomy', are still free to recognize such a group (for discussion, see Kitching et al. 1998: 12–13). Cladism seems to be replacing evolutionary taxonomy: accordingly, as I discuss other similar cases of 'correction' following, I will restrict my attention to cladism and ignore other schools like evolutionary taxonomy.

3. There is dissent with respect to both cases. With regard to 'fish', some professionals would retain naturalness by applying the term not only to sharks and salmon, but also to whales, cows, and ourselves (see, e.g., Dahn et al. 2007: 311). And some professionals would exclude birds from the 'dinosaur' camp (see LaPorte 2004: 88–9 for references), which forces them to drop 'dinosaur' from formal use since it concedes that the *dinosaurs* are not a natural group (at least in one key respect: see note 2). Of course, by dropping the term 'dinosaur' from formal scientific use in order to exclude the birds, we can honour the traditional claim, 'none of the flora and fauna alive today is an example of a dinosaur: dinosaurs are extinct'. But to exclude the birds crosses with other salient traditional claims, like 'the dinosaurs are a natural biological group including such and such paradigms and all their descendants'. Whether we include or exclude the birds, we seem to refine the former use of 'dinosaur' and it seems implausible to say that one way is right at the expense of the other.
4. This seems to be happening with respect to some cases: see note 2.
5. Hacking 2007 adds much by way of development to the case of 'jade'.
6. One such consequence (discussed in LaPorte 2004, ch. 5) is that we face anew Kuhnian worries about the progress of science. A position like mine would seem to be at a loss to account for the way that science advances 'by progress in understanding': that seems incompatible, as one author complains, with advancement 'by fiat or majority vote of a committee' that changes the meanings of key theoretical terms (Weintaub 2007: 184; Slater thoughtfully examines Weintaub's argument, in 'Pluto and the platypus'). It was once common to appeal to the causal theory of reference to dismiss meaning change and attending worries like the foregoing but I have raised the worries even while embracing the causal theory of reference (see LaPorte 2004: 5–7, 112–20). Other worries that revisit in a new guise include Quinean worries about necessity (LaPorte 2004, ch. 6).
7. A published version is not available at time of going to press, so I cite Bird (2007) by section, with permission.
8. This observation about 'water', a term from the vernacular, evidently does not carry over to terms like 'hydrogen' and 'oxygen': biological and physical properties do not have the same value for determining the application of terms coined within, and primarily used within, the field of chemistry. So the vagueness associated, at baptism, with terms like 'hydrogen' and 'oxygen' is less extensive than the vagueness associated with 'water'. It seems plausible to say that whatever vagueness has attended 'hydrogen' or 'oxygen' has left scientists little room for discretion in determining the term's *extension* after baptism (whether the *essence* of these entities has been articulated is another matter, which I will set aside for now: but see LaPorte 2004: 109–10): both Bird and Robin Hendry indicate as much by way of highly effective discussions replete with thoughtful examples. I quote from Hendry, who claims about me, mistakenly, that I

   > would see the term 'oxygène' in the mouth of Antoine Lavoisier, who introduced it, as referentially indeterminate. When Lavoisier intro-

> duced 'oxygène,' the samples he ostended would have been mixtures of the isotopes oxygen-16, oxygen-17 and oxygen-18, which have different atomic weights, but the same nuclear charge. Lavoisier was innocent, of course, of the distinction between atomic weight and nuclear charge. So LaPorte is quite right that we cannot attribute to him the intention that 'oxygène' should apply to any, or all, of oxygen's isotopes. But to leave it at this would be to ignore potentially relevant historical facts: Lavoisier was attempting to name an *element*, a substance which could explain (and indeed survive) chemical change. Since atomic weight (with respect to which isotopes differ) is a vanishingly insignificant determinant of chemical behaviour compared to nuclear charge (which they share), it is overwhelmingly plausible to interpret 'oxygène' in Lavoisier's mouth as applying to *any* isotope of oxygen. (Hendry 2004: 72–3; much the same argument is provided in Hendry 2006: 868–9)

So far, so good, except for the suggestion that I would reject this account concerning 'oxygen'; what I *would* reject is not *that* but rather the alleged *parity* that Hendry goes on to suggest between 'oxygen', on the one hand, and '*water*' on the other hand (Hendry 2004: 73; 2006: 869): in the same way, I would reject the parity suggested by Bird (2007: §3) between 'the names "carbon", "potassium", and "uranium" ', on the one hand, and 'water' on the other hand (cf., by contrast, Donnellan 1983, whose case concerning 'water' and other natural kind terms is in many ways similar to mine, but whose case *does* seem to be undermined by observations like those of Hendry and Bird: for discussion, see LaPorte 2004: 190, n. 9).

9. I have reinterpreted TAX slightly but in the intended spirit; my criticisms apply to the original version too.
10. Here is a salient difference between 'water' and 'hot': the vagueness of 'water' has been refined away along dimensions that I have been discussing. The vagueness of 'hot', by contrast, has not. Within *specific* contexts established in specific conversations, there is refinement and so more or less precision in the use of 'hot', to be sure; but it is not as if, for *general* contexts, some borderline cases of 'hot' have lost their borderline status after linguistic change that was prompted by empirical discoveries exposing the borderline status. The borderline status of drinks like those I have been talking about was all along understood: that this much vagueness in 'hot' should be tolerated has long been a settled matter. But the borderline application of 'water' to D_2O became clear only after scientific discovery, so a decision whether to tolerate the vagueness had to be made at that point: again, the vagueness of 'water' was *open texture*.
11. Although Bird does not question it, I have already discussed, in note 6 and the accompanying text, an assumption here that the failure of discovery is interesting when it attends interesting scientific statements (whether or not theoretical identity statements *qualify* as interesting scientific statements). I defend the claim that theoretical identity statements are especially interesting in section 2.2.

> I do not hold that scientific change is *always* marked by meaning change that would frustrate claims to discovery (to the contrary, see LaPorte 2004: 135–6), but I do hold that such meaning change *commonly* attends scientific change, especially scientific change leading to the affirmation of theoretical identity statements. A nicety that I can only gesture toward here and that Bird's forceful criticisms have prompted me to think more about is that the relevant meaning change could come before the first *reports* that a theoretical identity statement

is true, so that the reports really do or would mark discoveries about a subject matter discussed in the *recent* tradition, though the reports fail to mark discoveries about a subject matter discussed in the *prescientific* tradition and in the more recent tradition up until some point at which meaning refinement might take place: thus, 'Rodentia' might now be refined so as to allow for the discovery, in the future, that horses and ourselves are included (see the discussion of Phylocode, in section 2.1).

12. You may hear me as discussing $gold_2$, or $lead_2$, which we can certainly name and discuss even if our words 'gold' and 'lead' don't quite match up.
13. Just as scientists can be more or less explicit about endorsing essence-exposing theoretical identity statements (see the foregoing quotations from Ridley), they can be more or less explicit about refining. 'Only in very few cases', the scientists Joyce, Parham, and Gauthier (2004: 991) note, can 'traditional name usage be inferred unambiguously': this is especially so for names coined in pre-Darwinian times. Just how to circumscribe taxa is a matter earlier authors left 'to the inclinations of individual reviewers', the foregoing scientists maintain, so it is up to individual workers to settle on some refined use or other for venerable taxonomic terms (for a similarly explicit example from mineralogy, see Dietrich and Chamberlain 1989).
14. My response here might prompt a follow-up question (it has, on occasion): what sort of essence is designated by vague words before the rise of science, and how would such an essence compare with the sort of essence that is designated after the rise of science, when vagueness is refined away? One possible response, if you permit me to speak of fuzzy essences with borderline extensions, is this: the earlier-designated essence is fuzzy in ways the later one is not. This suggestion will displease some. That is okay: there are other ways to accommodate vagueness. For example, we might say that vague terms oscillate between a range of essences, and that with refinement, that range narrows as vagueness is honed away (or the range narrows in certain contexts: see the discussion of H_2O and D_2O in section 2.1). The right account of how to understand the essences corresponding to terms whose vagueness shifts over time will depend on the right account of vagueness, and I do not take a position with respect to that.
15. It might be tempting to try to blame my account of terms' refinement for the problems that vagueness presents; that would be a mistake. My account does not *introduce* vagueness and its problems. Kripke and Putnam, as well as everyone else, are forced to acknowledge vagueness: even 'Hesperus' is vague, so a statement of the corresponding essence will be too. I acknowledge *more* vagueness in natural language than Kripke and Putnam do and again, I acknowledge a *refinement* of terms' use over time to *reduce* vagueness; but since the problems attending vagueness are *theoretical* problems of how to make sense of it in principle, nothing I say about the extent of the phenomenon in practice, including changes over time in the extent of the phenomenon, burdens us with a cost that we could avoid paying by resisting my claims. We are stuck with vagueness and the problem of how to make sense of it regardless of whether we accept my claim that important terms display, to important effect, more vagueness at earlier times than at later times.

REFERENCES

Bird, A. (2007) 'A posteriori knowledge of natural kind essences: a defence', *Philosophical Topics*, 35 (forthcoming).

Cantino, P. D. and de Queiroz, K. (2007) 'Phylocode: international code of phylogenetic nomenclature'. Online draft version 4b. Available HTTP: <http://www.ohio.edu/phylocode/index.html>.

Dahn, R. D., Davis, M. C., Pappano, W. N., and Shubin, N. H. (2007) 'Sonic hedgehog function in chondrichthyan fins and the evolution of appendage patterning', *Nature,* 445: 311–14.

De Queiroz, K. (1994) 'Replacement of an essentialistic perspective on taxonomic definitions as exemplified by the definition of "Mammalia" ', *Systematic Biology*, 43: 497–510.

———. (1995) 'The definitions of species and clade names: a reply to Ghiselin', *Biology and Philosophy*, 10: 223–8.

Devitt, M. (2009) 'Biological realisms', in H. Dyke (ed.), *From Truth to Reality: New Essays in Logic and Metaphysics*, London: Routledge, 43–65.

Dietrich, R. V. and Chamberlain, S. (1989) 'Are cultured pearls mineral? The continuing evolution of the definition of *mineral*', *Rocks and Minerals*, 64: 386–92.

Donnellan, K. (1983) 'Kripke and Putnam on natural kind terms', in C. Ginet and S. Shoemakers (eds), *Knowledge and Mind*, New York: Oxford University Press, 84–104.

Dupré, J. (1986) 'Sex, gender, and essence', in P. French, T. Uehling, and H. Wettstein (eds), *Midwest Studies in Philosophy 11, Studies in Essentialism*, Minneapolis: University of Minnesota Press, 441–57.

———. (2004) 'Review of *Natural Kinds and Conceptual Change*', *Notre Dame Philosophical Reviews*. Online. Available HTTP: <http://ndpr.nd.edu/review.cfm?id=1439> (accessed 25 May 2009).

Enç, B. (1986) 'Essentialism without individual essences: causation, kinds, supervenience and restricted identities', in P. French, T. Uehling, and H. Wettstein (eds), *Midwest Studies in Philosophy 11: Studies in Essentialism*, Minneapolis: University of Minnesota Press, 403–26.

Ghiselin, M. (1995) 'Ostensive definitions of the names of species and clades', *Biology and Philosophy*, 10: 219–22.

Greville, C., Count de Bournon, and Oakley, C. (1798) 'On the corundum stone from Asia. By the Right Hon. Charles Greville, F. R. S.', *Philosophical Transactions of the Royal Society of London*, 88: 403–48.

Hacking, I. (2007) 'The contingencies of ambiguity', *Analysis*, 67: 269–77.

Hendry, R. F. (2004) 'Review of *Natural Kinds and Conceptual Change*', *Philosophical Writings*, 27: 71–4.

———. (2006) 'Elements, compounds, and other chemical kinds', *Philosophy of Science*, 73: 864–75.

Hennig, W. (1965) 'Phylogenetic systematics', in R. Smith and T. Mittler (eds), *Annual Review of Entomology*, Vol. 10, Palo Alto, CA: Annual Reviews, Inc., 97–116.

Hughes, R. W. (1990) *Corundum*. London: Butterworth-Heinemann.

Hull, D. (1998) 'Taxonomy', in E. Craig (ed.), *Routledge Encyclopedia of Philosophy*, Vol. 9, London: Routledge, 272–6.

Joyce, W. G., Parham, J. F., and Gauthier, J. A. (2004) 'Developing a protocol for the conversion of rank-based taxon names to phylogenetically defined clade names, as exemplified by turtles', *Journal of Paleontology*, 78: 989–1013.

Kitching, I. J., Forey, P. L., Humphries, C. J., and Williams, D. M. (1998) *Cladistics: The Theory and Practice of Parsimony Analysis*. Oxford: Oxford University Press.

Kripke, S. (1980) *Naming and Necessity*. Cambridge, MA: Harvard University Press.

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

Lewis, D. K. (1986) *On the Plurality of Worlds*. Oxford: Blackwell.
Nelson, J. S. (2006) *Fishes of the World*. 4th edn. New York: John Wiley & Sons.
Putnam, H. (1975) 'The meaning of "meaning" ', in his *Mind, Language and Reality*, New York: Cambridge University Press, 215–71.
Ridley, M. (1989) 'The cladistic solution to the species problem', *Biology and Philosophy*, 4: 1–16.
Slater, M. H. (unpublished) 'Pluto and the platypus: tale of an odd ball and an odd duck'. Unpublished manuscript.
'Those diverse diapsids: the reptiles (except turtles).' 1994–2009. University of California, Berkeley Museum of Paleontology. Online. Available HTTP: <http://www.ucmp.berkeley.edu/diapsids/index.html> (accessed 6 May 2009).
Weintaub, D. (2007) *Is Pluto a Planet?* Princeton, NJ: Princeton University Press.
Wood, A. E. (1957) 'What, if anything, is a rabbit?' *Evolution*, 11: 417–25.
Wyss, A. R. and Meng, J. (1996) 'Application of phylogenetic taxonomy to poorly resolved crown clades: a stem-modified node-based definition of Rodentia', *Systematic Biology*, 45: 559–68.

7 Discovering the Essences of Natural Kinds

Alexander Bird

1. INTRODUCTION

Following Kripke, Putnam, and others, many hold that natural kinds have essences, and that these essences may be discovered *a posteriori*. Joseph LaPorte (2004) very carefully, and in many respects convincingly, articulates an alternative view of what is occurring. Concentrating on theoretical identities such as 'water is H_2O', LaPorte argues that there is considerable vagueness in the use of kind terms, especially vernacular kind terms. This vagueness is a matter of *open texture*. For a kind term '*K*', some things will be determinately *K* and other things will be determinately not *K*. But there will be a boundary of things for which there is no determinate fact of the matter whether they are *K* or not. This means that there will be no determinate fact of the matter that '$K_1 = K_2$' for distinct kind terms 'K_1' and 'K_2'. According to LaPorte, when a natural kind identity is established as being determinately true, that is because scientists have made a *decision* to adopt the identity as true. In so doing, it will now be determined of items that were previously in the boundary (neither *K* nor not-*K*) whether they are *K* or not. For example, we now regard heavy water (deuterium oxide) as a subspecies of water; but scientists could have decided to exclude deuterium oxide from the extension of 'water'. So 'water is H_2O' is true in virtue of a decision. That truth, LaPorte emphasizes, is indeed a necessary truth, but it is not the discovery of some previously hidden essence. Rather, it is an empirically motivated *stipulation*.

While I believe that LaPorte's discussion of these issues furnishes us with many important insights, I will argue in this article for the following claims, in successive sections:

- There is rather less room for conceptual choice and stipulation than LaPorte supposes. His view is that when a stipulation is made that '*V* = *S*' where '*V*' is a vernacular term and '*S*' is a more precise scientific term, there is a precisification of the vernacular term '*V*'. Within the prior vagueness of '*V*' there is considerable room for choice as to how the term might be precisified, which is why the truth of '*V* = *S*' is a matter of decision. I argue that the vernacular concept is governed by rules of application that do not leave such scope for decision.

- There are many essential truths whose truth cannot be accounted for in the way that LaPorte suggests. We should not focus too much on identities, since we can discover many essentialist truths that are not identities.
- A particular set of cases that LaPorte's view does not accommodate includes identities of the form '$K_1 = K_2$' where 'K_1' is not a vernacular term, but is a scientific term without the high degree of vagueness attributed by LaPorte to vernacular terms. Nevertheless, for these cases '$K_1 = K_2$' does not express an analytic truth, because 'K_1' is introduced before later discoveries in science that allow for the articulation of the identity statement. I expand on certain examples, in particular chemical elements discovered in the nineteenth century before atomic structure was understood.

2. CONCEPTUAL CHANGE AND PRECISIFICATION

In this section I argue that there is less scope for decision concerning the application of our vernacular kind concepts than LaPorte suggests. (Some of the points made in this section are discussed in greater detail in my 2007.) It is important to note that LaPorte is saying more than simply that our natural kind concepts can change and that our vernacular terms can be stipulated to have new extensions in the light of developments in scientific knowledge. LaPorte's more specific claim is that such changes are in an important way consistent with the existing use of the vernacular concept. When we agree that $V = K_2$ we are precisifying the concept V. Let us call the older vernacular concept V_O and the newer precisified concept V_N. We may think of V_O as having a determinate extension (i.e. things that are clear cases of V_O), a determinate anti-extension (i.e. things that are clearly not cases of V_O), and a boundary (things that are both not clearly V_O and not clearly not V_O). The concept V_N includes the determinate extension of V_O. A precisification divides the boundary between the extension of V_O and its anti-extension. But nothing in the prior concept, V_O, determines where within the boundary a legitimate division should fall. Anywhere is permitted, and the extension of V_N may include all or none of the boundary of V_O.

For LaPorte's case to be made, it is necessary that the relevant examples satisfy three requirements:

1. V_O and V_N are both natural kind concepts.
2. Any stipulation made with regard to the extension of V_N really is a matter of precisifying V_O rather than conceptual shift. (By 'conceptual shift' I mean a conceptual change from C_O to C_N where some x was C_O but is not C_N, or was not C_O and is C_N. Precisification is not

conceptual shift, because it involves no change to the determinate extension or anti-extension, only to the boundary.)

3. The concept V_O does not itself determine the division of the boundary region between its extension and anti-extension. (If V_O did this, then V_O and V_N would not really be distinct concepts.)

In my opinion, a counter-view to that of LaPorte can argue that each of his representative cases fails to satisfy one or more of the conditions listed.

For example, LaPorte considers the case of 'water = H_2O' (where H_2O includes all isotopic variants). He reruns a twin-earth-style thought experiment. In this thought experiment scientists consider D_2O, which is the oxide of deuterium, the isotope of hydrogen with mass number 2, whose atoms thus contain one proton and one neutron. D_2O is poisonous. In the light of that and similar information, LaPorte argues that the scientists could reasonably have precisified the term 'water', so that its excludes D_2O, with the consequence that 'water = H_2O is false (but 'water = P_2O' is true, where 'P' denotes protium, the isotope of hydrogen with mass number 1). I argue that scientists were not free to make such a decision without violating the concept *water* in a way that does not count as a precisification. That is, the concept *water*$_O$ did determine that D_2O falls within its extension. Had scientists decided to exclude D_2O from the extension of *water*$_O$, then *water*$_N$ would not have be a precisification of *water*$_O$. The principle upon which I make this claim asserts that there is a division of linguistic labour among scientists, such that it is the job of a particular subset of scientists to determine the facts concerning particular sorts of natural kinds. Thus it is the job of biologists to determine the nature of and relationships between the various sorts of organism, while it is the role of chemists to determine the nature and identity of substances. Thus Linus Pauling tells us:

> The different kinds of matter are called *substances*. Chemistry is the science of substances—their structure, their properties, and the reactions that change them into other substances. (Pauling 1970: 1)

In the light of this, it will be chemical facts that determine the identity of substances. The chemical facts class D_2O with other kinds of H_2O. The structure of D_2O is the same as that of the other isotopic variants of H_2O, and the reactions it engages in are the same. Its qualitative chemical properties are also the same. D_2O differs chemically from the other isotopic variants as regards certain quantitative features, such as rates of reaction. Strictly, all isotopic variants of all chemical substances differ from one another in such quantitative ways, but the difference is much more marked in the case of D_2O. As a result of this marked difference in reaction rate, pure D_2O in place of water can be poisonous for many organisms. LaPorte regards this as one of the reasons why scientists in his twin-earth story would be willing to exclude D_2O from the extension of *water*$_N$. Another reason is that D_2O

can be used in the manufacture of fusion bombs in a way that the other isotopic variants cannot. Note that both these reasons come from outside chemistry. They are reasons, therefore, that are not pertinent to the science whose job it is to investigate the nature of and to classify water. Given the principle enunciated earlier, I regard *water*$_O$ as including D_2O within its extension and so see no distinction between *water*$_O$ and *water*$_N$.

I am not suggesting that LaPorte's twin-earth story is implausible or that his hypothetical scientists would be unreasonable. Undoubtedly our concepts do indeed change, which is to say that a term may first express one concept, then another, distinct but usually related concept. Rather, I am saying that their decision would be a decision not to precisify the concept *water* but to shift it, so as to exclude, on sensible pragmatic grounds, a subset of its earlier extension. Thus the actual history of the concept *water* violates requirement 3, in the sense that the concept *water*$_O$ already determined whether the apparent boundary region between determinate water and determinate non-water should be drawn. The concept *water*$_O$ just is the concept *water*$_N$. On the other hand, the hypothetical change envisaged by LaPorte violates requirement 2, since his story involves a conceptual shift.

The pragmatic concerns that might stimulate a conceptual shift are likely to be present in the use of vernacular terms for kinds, since those kinds play a significant part in our social, economic, and cultural lives. I regard LaPorte's very informative discussion of the history of the term 'jade' in this light. LaPorte points out (in contrast to almost all other philosophical discussions) that Chinese jade workers and experts were fully aware that a new jade-like substance, jadeite, that was being imported into China was different from their traditional jade, nephrite. Nonetheless, they took a decision to regard the new material as an instance of 'jade', along with their traditional nephritic jade. LaPorte take this to show that we cannot be confident of Putnam's judgment that faced with XYZ and the facts of its composition, we would deny it is water. Oscar and his friends might very well decide that it is reasonable to call XYZ 'water', in order to indicate to other Earthlings that it is safe to use just like Earth water, just as they might, in LaPorte's story, decide to exclude D_2O because of its toxicity. But it does not follow from the facts that such decisions might reasonably be made that such decisions are precisifications as opposed to conceptual shifts. And even if the decisions are precisifications, it needs to be argued that the terms in question are indeed natural kind terms. So in the case of jade, we might agree with LaPorte that 'jade' (or, rather, the Chinese term 'yü') had an open texture such that while nephrite was determinately jade, jadeite was initially in the boundary region, and furthermore that Chinese jade workers then decided to precisify the term by including jadeite. But why should we regard 'jade' as a natural kind term? The fact that the extension of a term '*T*' is the extension of a natural kind term does not make *T* a natural kind. The extension of 'humodo' includes all humans and dodos, and so has a current extension that is a natural kind, but 'humodo' is not a natural kind term.[1] So even if the (determinate) extension of *jade*$_O$ was precisely nephrite, we cannot conclude from that fact that 'jade' was then a natural kind term. In my view, if the term failed determinately to exclude a substance of very different composition, then that shows it did not name a natural substance, and so requirement 1 fails. Equally, if 'jade' did name a natural substance, then the decision by Chinese jade workers and connoisseurs to admit jadeite as a kind of jade amounts to a conceptual shift, from a natural kind term that excluded jadeite to a term that is not a natural kind term and which includes jadeite.

LaPorte (2004: 97) anticipates such a response, remarking that 'the claim that that is the right moral to draw seems ill-supported and motivated only by a desire to save a theory'. However, my motivation is that alluded to earlier, that chemists regard it as a necessary condition of being the same substance that two samples share the same or very similar composition and engage in the same reactions. Clearly *jade*$_N$ is not the concept of a natural kind of substance, since it violates this condition. If *jade*$_O$ had the open texture that LaPorte ascribes to it, then it would equally have violated this condition, since this condition determinately excludes samples of jadeite from being the same substance as samples of (traditional) jade, whereas the open texture view leaves that indeterminate. So either *jade*$_O$ was not a natural kind concept (violating requirement 1) or it was a natural kind concept, but the change to *jade*$_N$ is not a precisification (violating requirement 2). Either way, the change from *jade*$_O$ to *jade*$_N$ is not a precisification of a natural kind concept.[2] If *jade* never was a natural kind concept, that would not be surprising if the relevant experts to whom we defer in the division of linguistic labour in deciding what determines the extension of a concept are in this case not chemists but are jade workers and connoisseurs, whose position is rather closer to those who have to decide whether sparkling wine made outside of the champagne region of France may be called 'champagne' or cheese manufactured other than in Somerset is correctly designated 'cheddar'.

Our discussion suggests that our concepts might fail to be natural kind concepts, because certain practical concerns make some other kind of concept more useful. Nonetheless, when classifying natural objects, classifying them by their natural kinds is one compelling way to go. Often practical interests and the desire to classify things by their natural kinds will coincide. We want to classify ores by their natural kinds, because the same natural kind of ore will produce the same metals and other minerals. We classify the metals by their natural kinds, because the instances of the same natural kind of metal will possess the same properties, the properties that make the metals useful for a distinct purpose. But, as in the case of jade, there may be interests that pull away from the natural kind concept. Another kind of practical concern is that of maintaining as much of an existing pattern of usage as possible in the light of discovery. Because of their superficial similarity and our ignorance of their deeper differences, items of kind *X* and kind *Y* might both be regarded as being of kind *K*. When the deeper differences are discovered, there will be a tension between, on the one hand, maintaining '*K*' as a natural kind term, and, on the other hand, the inconvenience of suddenly denying that *Y*s are *K*s. As in LaPorte's discussion, a choice will have to be made, but this does not mean that the older concept K_O does not determine an extension or that the choice is between different precisifications of K_O. We can choose, if we wish, to make a conceptual shift. LaPorte discussed a number of interesting biological cases. Some zoologists think that guinea pigs' latest common ancestral population shared with all other creatures commonly classed as rodents is very early, and that this ancestral population is also an ancestral population of many non-rodent mammals, such as horses. We could go in three directions: (a) deny that guinea pigs are rodents; (b) accept guinea pigs as rodents and regard the kind of rodents as including guinea pigs and all the other standard rodents, *plus* the intermediate kinds (horses etc.); *rodent* would still be a natural kind concept; (c) accept guinea pigs as rodents and regard the kind of rodents as including guinea pigs and all the other standard rodents, but *excluding* horses and so forth; 'rodent' would be a polyphyletic classification—covering two kinds but not the kinds in between, and as such $rodent_N$ would not be a natural kind concept. While (a) is the choice of some zoologists, LaPorte points out that (b) and (c) have precedents. When it was discovered that birds are descended from dinosaurs, it became widely accepted that birds are living dinosaurs. On the other hand, that fact also means that if *reptile* were to be monophyletic (covering all of just one clade), then birds would have to be classed as reptiles. A consequence of this has been a move to exclude the term 'reptile' from taxonomy, because it is paraphyletic (covering a clade minus one subclade, the birds).

The fact that different choices might be made may suggest (as LaPorte holds) that there is open texture that permits these different precisifications of a natural kind term such as 'rodent'. While I should repeat that a choice does not imply a precisification (it could be a shift), there is, however, a pattern to the decisions made that suggests that our past uses of the relevant kind terms does determine an extension. The idea is this. We use the relevant terms with the intention that they name a natural kind. Which kind, *K*, is that? The answer is given by the principle (TAX): (i) *K* must be a clade (i.e. '*K*' should be monophyletic—which is to say, it really does pick out a natural kind); (ii) a clear majority of subtaxa regarded as paradigmatic of *K* should be included in the extension of *K*. (iii) a clear majority of subtaxa regarded as typical foils for *K* (i.e. paradigmatic non-*K*s) should be excluded from the taxon. If it is impossible to meet these requirements, then '*K*' does not pick out a natural kind, in which case it may well be natural to continue to use it in the vernacular, as a polyphyletic or paraphyletic non-natural kind term. While there may well be residual vagueness, it does determine answers to 'are guinea pigs rodents?' (no: because rodents form a clade by excluding guinea pigs but not mice, rats, gerbils, and so on, and without including horses and primates), and 'are birds dinosaurs?' (yes: because dinosaurs taken to include birds form a clade; birds are not important, paradigmatic foils for dinosaurs, since the paradigmatic foils for dinosaurs are not modern creatures but are Mesozoic animals such as the Triassic Crocodylia and the thecodonts). Lastly, if we ask 'are birds reptiles?' the answer 'yes' would include as reptiles an important group of paradigmatic foils for reptiles, the birds. On the other hand, the answer 'no' makes it impossible to regard reptiles as forming a clade. 'Reptile' must therefore be excluded from scientific taxonomy, and not be regarded as a natural kind classification. But it may be retained in the vernacular as a paraphyletic, non-natural classification.

In this discussion I have not disproved LaPorte's claim that natural kind terms have an open texture that allows for precisification. Rather, I have contrasted his view with another which also fits the data in a non-arbitrary and, I hope, non-question-begging way, and which regards natural kind terms as having much more determinate extensions.

3. ESSENTIAL TRUTHS WITHOUT ESSENCES

In this brief section I argue that LaPorte's arguments about identities, and the claim their truth is stipulated, leave many essentialist claims untouched. I shall not press this point at length because LaPorte accepts it. Nonetheless, it is worth reiterating because it illustrates two distinct forms of argument for essentialism. One draws its force from Kripke's discussion of names and rigidity. If '$K_1 = K_2$' expresses a true proposition, and if 'K_1' and 'K_2' are rigid designators, then it is necessarily the case that $K_1 = K_2$ (the rigid designators might be rigidified definite descriptions). Furthermore, depending on the content of the term 'K_2' '$K_1 = K_2$' *might* reveal the essence of 'K_1', as in 'water = dthat (the compound H_2O)' (although there is of course no guarantee—some further argument is required in any particular case to show that the necessity thus produced is an essentialist necessity). However, according to LaPorte, it might not be determinate which natural kind 'water' refers to, and so there is room for different precisifications, one of which makes 'water = H_2O' true, others of which do not. If we choose to precisify 'water' in that way, then 'water = H_2O' comes out as necessarily true, for the reasons given. But this necessity is the result of stipulation.

On the other hand, we may establish essentialist claims via a different route, such as Putnam's twin-earth thought experiments, and the intuitions to which Kripke appeals when discussing, for example, the essentiality of origin. Such arguments do not appeal to the necessity of identity or to the properties of rigid designators. Rather, they appeal to intuition, for example, the intuition that if something is XYZ rather than H_2O, then it would not be water. Such arguments do not of themselves establish identities. For the intuition referred to gives us: necessarily, if *S* is water, then

S is composed of H_2O. But it does not yield: necessarily, if *S* is composed of H_2O, then *S* is water. That would require an additional argument, and is harder to establish. (Indeed, Putnam seems to imply that this is false, because he, perhaps carelessly, implies that water is a liquid, so solid or gaseous H_2O is not water.) As a result, such claims are less sensitive to the open texture that LaPorte holds to exist, even if he is correct. The open texture of 'water' may leave it open whether it designates the kind H_2O or the kind P_2O, which is itself a subkind of H_2O. But either way it will be (necessarily) true that if *S* is water, then *S* is composed of H_2O.

Another illustration of the fact that open texture does not exclude all essentialist claims concerns the classification of animals that *clearly* fall within a clade. A newly discovered species might fall determinately within a clade, even if the boundaries of that clade are vague. So LaPorte points out that the genealogy of the giant panda puts it somewhere between the brown bear (a bear) and a racoon (a non-bear). He says that it is a matter of decision whether the bear clade should include pandas or not. Let us allow that LaPorte is right—and if one believes in any open texture at all, this would seem to be a plausible case. But that fact does not mean that it is similarly a matter of decision, undetermined by the concept *bear*, whether to classify as a bear the rare Tibetan blue bear, first considered by European naturalists in 1854. For this new discovery is a subspecies of the brown bear. So *every* precisification of 'bear' includes the Tibetan blue bear within its extension. So the Tibetan blue bear is essentially a bear, and that fact is a matter of discovery (since its ancestry is a matter of discovery).

4. DISCOVERING ESSENCES OF THEORETICAL KINDS

Finally I will argue that LaPorte ignores a range of important cases where essentialist claims, including identities, can be established without facing the problems that he identifies. LaPorte concentrates largely on claims involving vernacular kind terms. These help LaPorte make his case, since one might suppose that open texture is going to be more likely and broader for terms whose use is established without the benefit of scientific theory. Furthermore, the fact that their use is more susceptible to influence by extra-scientific concerns makes their extensions appear subject to choice and decision.

It might appear that the traditional essentialist, for whom essences and identities are discovered not stipulated, faces a dilemma. Sentences (such as identities) employing vernacular terms can be treated using LaPorte's analysis. On the other hand, sentences employing technical, scientific terms look as of they too may be regarded as stipulations, even as analytic. Take 'mendelevium = the element with atomic number 101'. The team at Berkeley which in 1955 synthesized mendelevium by bombarding einsteinium with alpha particles knew precisely that they were synthesizing the element with atomic number 101, and on confirming that they had done so, proposed the name mendelevium for the new element. In this case it looks very much as if 'mendelevium = the element with atomic number 101' is analytic, certainly at least that this is a stipulation.

The response to this worry is to show that this is a false dilemma. It is false because many natural kind terms are scientific, not vernacular terms, yet they were introduced in advance of the relevant discoveries that would permit making a stipulative or analytic definition in a manner analogous to 'mendelevium = the element with atomic number 101'. The key point in what follows is that a certain amount of scientific knowledge is required to introduce a new natural kind term, knowing that it is indeed a kind term and knowing that it names a different kind from other kind terms similarly introduced. However, this amount of knowledge can be *less* than the amount of knowledge required to know the essences of those kinds. Hence when essence-stating identities do become known, they must be known *a posteriori*.

We can articulate this point with respect to the chemical elements as follows: it was possible to know that a newly introduced name for an element (e.g. 'actinium', circa 1899–1904) did indeed name a distinct element. But it was not possible until later in the twentieth century to know that 'actinium = the element with atomic number 89'. Hence the latter proposition cannot be a stipulation or analytic.

Nine elements were known in ancient times. Arsenic, antimony, and zinc were identified as distinct substances in the medieval and renaissance periods. Phosphorus is the first element to be isolated by what we might think of as modern, chemical means, by Hennig Brand in 1669. The rise of modern chemistry in the eighteenth century saw the identification of several new elements: bismuth, platinum, nickel, and cobalt were identified between 1732 and 1753. The chemical revolution centred on the investigation of gases, and the discovery of hydrogen, oxygen, nitrogen, and chlorine between 1766 and 1774.

The important theoretical development was the introduction, primarily by Lavoisier in his *Traité Elémentaire de Chimie*, of the idea of a chemical element. Although the idea of an element goes back to Plato, Empedocles, and Aristotle, Lavoisier (1789: xvii) gave an account of chemical element as that which cannot be decomposed further by chemical analysis: 'we associate with the name of elements, or of the principles of substances, the idea of the furthest stage to which analysis can reach'. On this basis, Lavoisier listed 33 elements rather than the traditional four or five (earth, air, fire, water, and quintessence). While fire remains in Lavoisier's list as caloric, Lavoisier lists several different kinds of elemental 'airs'. Water, on the other hand, is a compound, since it decomposes into oxygen and hydrogen, as Lavoisier showed by passing steam over heated iron. The chemical revolution thus instituted a new paradigm of normal chemical science—isolating new chemical elements.

The next theoretical advance of significance was Mendeleev's development of the periodic table. As is well known, Mendeleev noted a correlation between atomic weights and certain chemical properties which repeated in a periodic fashion. The ordering of the elements by atomic weight thus gave rise to a new chemical property, the position of an element in the periodic table, its atomic number. This number was not simply the ordinal number of the element in the list of known elements ordered by empirically measured atomic weight, since Mendeleev used the periodic property to identify gaps in the table to be filled by hitherto undiscovered elements and in some case reversed the ordering of atomic weights. It thus may appear that Mendeleev's work identifies atomic number as the crucial identifying property of elements. On the other hand, it seems clear that Mendeleev did regard atomic *weight* as fundamental, for in the case of the reversed ordering (where the element with lower atomic number has a higher recorded atomic weight), Mendeleev thought that the atomic weight would need to be corrected. The order of *correct* atomic weights would be the order of atomic number.

So although Mendeleev introduced the notion of atomic number, this did not have a role distinct from that of atomic weight, which remained the fundamental notion. This was overturned in 1914 by Henry Moseley, who showed that there is a relationship between atomic number and X-ray frequency. The square root of the latter is proportional to nuclear charge in the Rutherford–Bohr model of the atom. Hence Moseley was able to show that atomic number is exactly equal to nuclear charge. On this basis Moseley could demonstrate that nickel's atomic number is one greater than cobalt's without having to maintain that the measured atomic weights (cobalt's atomic weight being greater than nickel's) needed correcting.

When we consider the sentence 'actinium = the element with atomic number 89', it is only once the notion of atomic number is associated with atomic charge, rather than position in a table ordered by atomic weight and periodicity of properties, that we can regard the sentence articulating the essence of actinium. For it is Moseley's notion of atomic number that is the explanatory one, which tells us the nature of the element.

The period between Lavoisier and Mendeleev was a period when scientists were able to isolate, identify, and distinguish new elements as such, but were not in position to know the truth of sentences of the form '*X* = the element with atomic number *N*'. During that period, 39 elements were discovered, in addition to the ten elements isolated during the eighteenth century before 1789.[3] To these we should add the 22 elements discovered after Mendeleev and before Moseley's work, since, although during that period sentences of the form '*X* = the element with atomic number *N*' were known, they did not express essential truths. Either way, we have a large number of elements discovered, which were given scientific names. That is to say, there is no question but that the names were used to name those natural kinds. Furthermore, they named exactly the same natural kinds as we refer to using the same names. There has been no conceptual change as regards these names of elements, not even precisification. There propositions of the form 'X = the element with atomic number N' concerning the 61 elements in question are propositions stating the essences of natural kinds that are known *a posteriori* rather than stipulated.

5. CONCLUSION

Whether any concepts have open texture is itself a debated question. However, on the assumption that there are such concepts, LaPorte makes a strong case for vernacular natural kind terms having open texture. That, he argues, leads to the truth of essence-stating identities being a matter of decision and stipulation rather than discovery. In the paper I have argued for three points:

(i) There is an alternative account of the data about what we do or would say in the light of new information. LaPorte argues that the data show that there is room for choice, which is explained by the open texture of the concepts. I respond that it is difficult to distinguish the precisification of an open-textured concept from a conceptual shift (involving a reassignment of extension). The data might equally be interpreted as the conceptual shifting of concept with little or no open texture, where the shift may be motivated by extra-scientific concerns. What seem to be instances of choices being made in varying directions can be explained by an account, (TAX), of the relationship between our concepts and paradigm instances and foils.

(ii) Even if we accept the case for open texture, that leaves many propositions concerning essential properties to be discovered *a posteriori*. Vagueness between red and orange leaves it determinate nonetheless that a ripe tomato is determinately red. Likewise, while it may be indeterminate whether $K_1 = K_2$, it may be perfectly determinate that the kind K_3 is a subkind of K_1. And so it may be an essential property of K_3 that K_3s are also K_1s. Thus the concept *water* may have open texture so that it is not determinate whether D_2O is water. But that is consistent with its being determinate that all water is H_2O. It may have been indeterminate whether the giant panda is a bear, but fully determinate that the Tibetan blue bear is a bear.

(iii) The claims about open texture are most plausible concerning vernacular terms that appear to be natural kind terms. But important identity statements in science are not just those that conjoin a vernacular kind term with a scientific one. They include those joining one scientific expression with another, as in 'actinium = the element with atomic number 89'. Such statements are necessary truths (indeed, essentialist truths), but are discovered *a posteriori*. That is possible because is it

> possible to have knowledge of the determinate identity and difference of kinds without knowledge of the truth of such essentialist facts. In the case of the chemical elements, the techniques of laboratory analysis were sufficient for chemists to isolate, describe, and identify new elements, and so name them, before any knowledge of atomic number was introduced.

Thus, I conclude, LaPorte's persuasive and informative arguments notwithstanding, that there remains considerable scope for essentialists to maintain that many essentialist truths, both those asserting essentially necessary conditions and those asserting identities, are not stipulated but are discovered *a posteriori*.

NOTES

1. It might be objected that the extension of 'humodo' includes not only current humans and dodos but also past and future ones, and so does not have a natural kind extension. We can change the example, then, to 'humalien' whose extension includes humans and bug-eyed aliens from Mars (presumably nonexistent at all times).
2. The concept $jade_N$ clearly violates requirement 1. But that is not relevant here, since LaPorte's point about jade is to illustrate the fact that natural kind terms can have open texture and so natural kind concepts do not determine the response that Putnam claims they do. It is consistent with that view that some decisions about how to precisify an open-textured natural kind concept should lead to a non-natural kind concept.
3. The numbers are to be regarded as inexact, since the discovery of elements in the eighteenth century and, in some cases, in the early nineteenth century is a vague matter.

REFERENCES

Bird, A. (2007) 'A posteriori knowledge of natural kind essences: a defence', *Philosophical Topics*, 35 (forthcoming).

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

Lavoisier, A. (1789) *Traité Elémentaire de Chimie*. Paris: Cuchet.

Pauling, L. (1970) *General Chemistry: An Introduction to Descriptive Chemistry and Modern Chemical Theory* (3rd edn). San Francisco: W. H. Freeman.

8 The Elements and Conceptual Change

Robin Findlay Hendry

1. WHAT CHEMICAL KINDS HAVE TO DO

Antoine Lavoisier introduced the term 'oxygen' into chemistry,[1] but how is it possible that he could think and speak about the substance whose name he gave us? The problem is that modern chemistry individuates elements by their nuclear charge (equivalently, the number of protons in the nucleus), while Lavoisier was beheaded long before there was any widespread acceptance of the idea that chemical elements are individuated by types of atoms, and he was entirely innocent of the notion of nuclear charge. The issue arises for familiar reasons. Chemists have radically changed their opinions about the elements, so unless the subject matter of their thought is in some way independent of their opinions about it, it is hard to see them as making a series of discoveries about a constant subject of research, namely oxygen. Oughtn't we to see the chemists instead as a series of scientific Humpty-Dumpties, each defining a new topic for discussion? Not only that: we would like to say that Lavoisier thought erroneously that oxygen is a constituent of all acids, and that gaseous oxygen is a compound of oxygen and caloric, the substance of heat. But unless Lavoisier was able to think and speak about oxygen, it is hard to see how he can be said to have had any erroneous opinions about it.

In the 1970s, scientific realists like Hilary Putnam and Richard Boyd addressed precisely this kind of question. Putnam's account of the meaning of theoretical terms told us that Lavoisier was able to refer to oxygen because he had access to samples of it, or was able to formulate descriptions that, in Lavoisier's context, picked out samples of oxygen, and the new name he introduced referred to anything that was the same as his samples. But the same in what sense? Lavoisier's samples would have been similar in various ways to many different kinds of stuff. This issue, which has come to be known as the 'qua' problem, has been the subject of much discussion, the general conclusion being that Putnam's model of reference needs significant revision.

Further discussion of these familiar topics may seem an unedifying prospect, yet it is worthwhile in order to emphasize a number of themes. First, the widespread phrase 'Kripke-Putnam semantics' neglects the quite distinct problem situations that motivated Kripke and Putnam.[2] Putnam's was a semantic project in the philosophy of science, giving a model of the meaning of theoretical terms according to which continuity of reference is possible, to set against the consensus view that had emerged from positivism, whose consequences (semantic incommensurability) were made explicit by Kuhn and Feyerabend (see Putnam 1975: 235–8). Kripke's proposals, though of course intimately connected in their semantic substance, were more ambitious in their motivation, hoping to establish the existence of necessary *a posteriori* truth on the basis of a semantic theory for proper names and its analogue for natural kind terms (Kripke 1980). These are important differences because I believe that something like Putnam's account of reference can be used to do the job *it* was intended for without any commitment to Kripke's apparatus of identity statements, modality, and rigid designation, or his arguments for essentialism (though it was an elegant apparatus that Putnam borrowed in many of his own discussions). I will seek to demonstrate this possibility in sections 2, 3, and 4 of this paper, arguing that it is plausible to regard nuclear charge as the quantity that determined the extensions of the names of the elements in the eighteenth and nineteenth centuries, long before its discovery in the twentieth century. I will address the issues of necessity and essence only in section 5.

A second theme is that Putnam and his critics have for the most part pursued their arguments by constructing thought experiments, carefully designed to elicit semantic intuitions of one kind or another. But these are 'just-so' stories, constructed under so few constraints that it is unclear what significance they have for how real words work. In some cases it is not even clear that the envisaged scenarios are genuinely possible. Few people take the apparently simple course of referring their arguments to the *actual* history of science: in the case of the element names, I argue that this offers plausible ways to fill in the well-known lacunae in Putnam's model.

A third theme concerns the classificatory pluralism, and the contingency of science. Chemical classification has been shaped by its (historically contingent) epistemic interests, just as it has been shaped by the forces that govern chemical change. Consequently, chemical classification might well have developed differently had those interests been different. But one can agree that there are many different kinds of similarity and difference in nature that may form the basis of systematic scientific study, and that it is contingent that scientists study the ones that they do. This does not, however, mean that chemical categories are constructed, or projected onto a nature that is somehow devoid of them 'in itself'. Quite the opposite: the recognition that there are many divisions in nature excludes that there are none, and the recognition that a particular discipline might have studied *different* divisions has no tendency to undermine the reality of the divisions it does study, or to establish that there is no fact of the matter about which they are (see Boyd 1990: §3). Nor do such possibilities diminish in any way the sense in which those divisions can be said to have been discovered. That is why John Dupré's promiscuous realism is a viable position (Dupré 1993): the only respect in which the view of chemical classification I set out here qualifies Dupré's position on biological species is that, so far as I can see, chemistry has been more unified in respect of its classificatory and explanatory interests than are the biological sciences.

2. LAVOISIER ON THE ELEMENTS

How, then, could one argue that Lavoisier was able to refer to oxygen? That requires some detail of the context in which Lavoisier used the term: the material context (the substances he and others were able to collect), and also the intellectual context (how he reasoned about those substances). Turning briefly to the material context, Lavoisier gave detailed accounts of his experiments in a number of memoires, and in his *Traité Élémentaire de Chimie* (1790). If those chemical operations were governed by the same laws of nature as are in force now, there can be little doubt that Lavoisier and his contemporaries were able to collect samples of a number of gases, including oxygen and hydrogen. Of course these would have been impure samples, but (taking oxygen as the example) it was the oxygen present that was responsible for the chemical behaviour he documented. On Putnam's account, therefore, Lavoisier would have been able to introduce a kind term as referring to chemical substances that were the same as his samples. But the same in what sense? That brings us to the more involved question of the intellectual context: how did Lavoisier reason about elements?

Lavoisier commented on the notion of element in a well-known passage in the preface to the *Traité*:

> [I]f we apply the term *elements*, or *principles of bodies*, to express our idea of the last point which analysis is capable of reaching, we must admit, as elements, all the substances into which we are capable, by any means, to reduce bodies by decomposition. Not that we are entitled to affirm, that these substances we consider as simple may not be compounded of two, or even of a greater number of principles; but, since these principles cannot be separated, or rather since we have not hitherto discovered the means of separating them, they act with regard to us as simple substances, and we ought never to suppose them compounded until experiment and observation has proved them to be so. (1790: xxiv)

Call this an analytical approach to the elements (following Cassebaum and Kauffman 1976). It has sometimes been presented as central to the achievement of the chemical revolution, concentrating chemical minds on concrete laboratory substances, and ruling out the kind of *a priori* metaphysical speculation about the ultimate components of substances that was so common before the eighteenth century. There is some truth in that: Lavoisier presented his version of the analytical conception in just those terms (1790: xxii–iv), and *a priori* theories about the ultimate components of substances did peter out during the eighteenth century. However, it was far from original to Lavoisier, and can be found in the writings of Robert Boyle (1661: 350), and even in Aristotle, whose conception of the elements is quite opposed to Lavoisier's (Needham 2009b). That should give some pause for thought before identifying it as the modern chemical idea that ushered in the chemical revolution. I would also resist it being characterized as an 'empirical definition,' because it cannot be applied without theoretical interpretation. It is at best a criterion for when some substance should be counted as an element, rather than an account of what 'element' means. When Lavoisier gives detailed descriptions of chemical substances, their components, and the changes they undergo, he employs a deeper explanatory notion of element.

There are two reasons for denying that Lavoisier's quote expresses a simple empirical criterion: one is quite general, the other specific. The general reason is as follows. Suppose that it is established by simple observation that a chemical change of some kind takes place when some reagents are brought together. A decomposition is a reaction in which one of the reagents is reduced to its components, that is, substances which are compositionally simpler. Hence the conclusion that decomposition has occurred requires some conception of the relative compositional simplicity of the substances involved in the reaction. But that is just the sort of thing for which we are seeking an empirical criterion. The analytical criterion cannot, therefore, be applied on the basis of empirical information alone, but only in the context of broader chemical theories concerning which (classes of) substances are composed of which others. This is not just a logical point: it was a live issue in Lavoisier's time, for the relative compositional simplicity of whole classes of substances was controversial in the late eighteenth century. That was the substance of Lavoisier's debate with the phlogistonists. For instance, Lavoisier saw calxes as compounds of a metal and oxygen, while the phlogistonists saw metals as compounds of a calx and phlogiston. Correspondingly, the phlogistonists saw calcinations (i.e. the formation of a calx) as decomposition of the metal, while Lavoisier saw it as involving the decomposition of oxygen gas, which he thought to be a compound of 'base of oxygen' and caloric. Hence application of the analytical criterion cannot have been simply empirical because it relied upon interpretations informed by controversial compositional theories. From a historiographical point of view, a body of information can be simply empirical only if it is grounded in observation and its interpretation is uncontroversial.[3]

Specific examples of Lavoisier's application of the notion of 'simple substance' directly illustrate the more general point: he allows his compositional theories about acids, gases, and earths to dictate compound status for substances which he does not claim to have decompounded. For instance, he says of the acid of sea salt (hydrochloric acid, HCl):

> Although we have not yet been able, either to compose or to decompound this acid of sea-salt, we cannot have the smallest doubt that it, like all other acids, is composed by the union of oxygen with an acidifiable base. We have therefore called this unknown substance the *muriatic base*, or *muriatic radical*. (1790: 72)

Sure enough, it is the 'muriatic radical' that appears in Lavoisier's table of simple substances (1790: 175), rather than the as yet undecompounded acid of sea salt. Why did he not have the 'smallest doubt' about its composition? In Chapter V of the *Traité*, Lavoisier had shown how sulphur, phosphorus, and carbon all combine with oxygen from the air to form compounds which, when dissolved in water, display characteristically acidic behaviour (the previous quote about the acid of sea salt appears in Chapter VI). He infers from this a general claim about the composition of acids, that 'oxygen is an element common to them all, which constitutes their acidity; and that they differ from each other, according to the nature of the oxygenated or acidified substance' (1790: 65). This use of what might be called 'chemical analogy' to reach conclusions about the composition of substances is not an isolated case. On similar grounds he regarded gases as compounds, with caloric their common component (Lavoisier 1790: Ch. I; see also Hendry 2005); he also declined to list potash and soda on his table of simple substances because they are 'evidently compound substances, though we are ignorant as yet what are the elements they are composed of' (1790: 178). It might be objected here that Lavoisier's theory of acidity was, like his other compositional theories, based on empirical information, so that in some broad sense his justification for regarding acid of sea salt as a compound is empirical. That is true, but Lavoisier's compositional theories are not (of course) simple empirical generalisations but hypotheses accepted for their ability to explain a range of chemical phenomena. So the central point about the analytical criterion must be conceded: it is at best a guide that can be applied only within a framework of compositional hypotheses that must answer to overall explanatory concerns.

Now Lavoisier often appealed to changes in the weights of substances in drawing compositional conclusions, for instance, the fact that when a metal transforms into a calx in a sealed container, its weight increases in line with the reduction in weight of the air in that container (1790: 78), or the fact that when hydrogen and oxygen are burned together, the weight of water produced is equal to the combined weights of the hydrogen and oxygen used up (1790: Ch. VIII). This might be taken to imply some condition on relative compositional simplicity. Thus, for instance, if the substances X (e.g. a metal) and Y (some component of the air) react together in a sealed container producing substance Z (e.g. a calx), and the combined weight of X and Y is equal to that of Z, then X and Y are components of Z. But application of such a criterion requires theoretical assumptions about what can, and cannot, pass freely through the walls of the container. As

Paul Needham has pointed out (2008: 70), it is also complicated by the widespread supposition in the late eighteenth century, common to both Lavoisier and the phlogistonists, that some of the key agents of chemical change are imponderable (that is, weightless): Lavoisier himself listed both light and caloric in his table of simple substances; as we have already seen, he regarded the component of air that reacts with metals (i.e. oxygen gas) as a compound of oxygen and caloric. This means that oxygen *gas* is not a component of water, or of metal calxes. Rather one of its components is: the base of oxygen gas. Since caloric is assumed to be weightless, he could not have used gravimetric measurements to establish that it is a component of oxygen.[4] So weight relationships, though empirically available, cannot be regarded as determining compositional relationships. One last consideration is, I think, conclusive. Knowing that the weights of elements *X* and *Y* are equal to the weight of *Z* tells us nothing about the composition of *Z* unless we assume that weights are conserved across chemical change. That assumption makes perfect sense if we assume that *X* and *Y* themselves survive chemical change. Otherwise its justification is a mystery. That brings us to Lavoisier's deeper explanatory conception of an element.

The analytical criterion isn't a definition of 'element' because it cannot justify the assumptions that Lavoisier makes about elements when describing substances, their components, and the chemical changes they undergo. These assumptions, which I have elsewhere called the 'core conception' of a chemical element (Hendry 2006), embody the compositional idea that elements are the building blocks from which substances are formed. Lavoisier's key assumptions are as follows: (i) elements survive chemical change; (ii) compounds are composed of them; and (iii) the elemental composition of a compound at least partly explains its behaviour. Applications of these assumptions abound in the compositional theories and inferences we have already seen: oxygen is present in all acids, and so must survive the chemical changes that form them; caloric is present in all gases, and so must survive the chemical changes that form them; the chemical behaviour of a substance is a guide to its elemental composition, and so elemental composition must be involved in the explanation of chemical behaviour. Moreover, his descriptions of chemical processes assume elements to be the building blocks of substances, their continued existence immune to chemical change. Consider, for instance, the following account of the combustion of metals, from the preface to the *Traité*:

> Metallic substances which have been exposed to the joint action of air and fire, lose their metallic lustre, increase in weight, and assume an earthy appearance. In this state, like the acids, they are compounded of a principle which is common to all, and one which is peculiar to each. In the same way, therefore, we have thought proper to class them under a generic name, derived from the common principle; for which purpose, we adopted the term *oxyd*; and we distinguish them from each other by the particular name of the metal to which each belongs. (Lavoisier 1790: xxviii)

He also applies the names of particular elements without regard to their state of chemical combination. On the subject of metals again, he describes how they are found in nature:

> The metals, except gold, and sometimes silver, are rarely found in the mineral kingdom in their metallic state, being usually less or more saturated with oxygen, or combined with sulphur, arsenic, sulphuric acid, muriatic acid, carbonic acid, or phosphoric acid. (1790: 159)

The analytical criterion cannot justify thinking of elements in this way because it suggests chemical operations whose product is a particular laboratory substance, something you can have in a jar (what Paneth 1962 calls a 'simple substance'). Take, for instance, iron: the simple substance is a metallic solid at room temperature. What happens when iron reacts with air to form a calx? Does the iron persist? Judging by appearances alone the iron disappears, being used up (that, in fact, is the phlogistonist interpretation). The analytical criterion provides no different answer because it is silent on whether the products of analysis are actually present in the compound. But if we think of elements not as simple substances but as material *components* of substances that can survive chemical change (what Paneth 1962 calls 'basic substances'), then the iron does persist, for it is a common component of the metallic iron and its oxide. Once we attribute to Lavoisier the notion of a basic substance, simple substances can be understood in their turn as laboratory substances that contain only one such component. Thought of in this way, which is the only way in which simple substances have a clear connection to compositional theory, the compositional notion of a basic substance is prior to that of a simple substance. Lavoisier's theoretical project is then revealed as one in which he understands the behaviour of composite substances as the result of the basic substances they contain. Since the surest experimental demonstration of the composition of a substance is its resolution into simple substances corresponding to the basic substances of which it is composed, the analytical criterion is also recovered, as a derived condition. But we should not forget that the very application of the notion of a simple substance presupposes a network of compositional hypotheses that already employ the compositional idea of elements as substance components.

3. ELEMENTS BEFORE AND AFTER LAVOISIER

3.1 Before Lavoisier

Aren't Lavoisier's assumptions about the elements just obvious? No: in assuming that elements survive in compounds, Lavoisier stands in opposition to Aristotle's theory of chemical combination, according to which elements are *not* actually present when combined in a mixt.[5] Aristotle's view was developed in conscious opposition to atomism, according to which the ultimate components of things persist unchanged in more complex bodies, the differences between things being explained by their different arrangements. Aristotle argued that if elements combine to form a new substance, as opposed to merely being juxtaposed, the product must be homogeneous. However, atomism can accommodate only juxtaposition, and so cannot recognise this distinction between mere juxtaposition and genuine combination (Needham 2009b). Aristotle's positive counterproposal is to generate the elements from opposed pairs of properties: hot and cold, wet and dry. The elements correspond to maximal degrees of compatible pairs of these properties: air is what is hot and wet, water is what is cold and wet, fire is what is hot and dry, and earth is what is cold and dry. In combination the essential properties of the elements are blended, so that a mixt will have submaximal degrees of heat or cold, wetness or dryness. Since a mixt is homogeneous, the elements are not actually present in the mixt because no part of the mixt possesses the essential properties of any of the elements. What does persist? Needham interprets Aristotle's view as taking the continuants of change to be bodies of matter that bear substance properties, including a potential to display the elemental properties once again (2009b: 9–10). Only in this sense—potentiality—can the elements be said to be 'in' the mixt.

Conceptions of elements (or principles) found in chemical texts of the seventeenth century stand in striking contrast to Lavoisier's (Siegfried 2002: Ch. 1; Klein and Lefèvre 2007: Ch. 2). In fact they are much closer to Aristotle's view, though they are much less clear and systematically developed. The elements are few in number (whether 3, 4, or 5) and they are identified *a priori*. The resulting compositional theories have only a distant relationship to the understanding of particular kinds of chemical change studied in the laboratory. Elements are not viewed as material parts of laboratory substances, but instead contribute to the character of a composite substance by offering their characteristic properties: as in Aristotle, the properties of a composite are a blend of those of the elements. It might be thought that this kind of speculation about the elements was swept away by healthy doses of empiricism and mechanical philosophy. Robert Boyle tried to administer both. In *The Sceptical Chymist* he clearly conceives of composition as material, and sets out the analytical conception of the elements:

> And, to prevent mistakes, I must advertize you, that I now mean by Elements, as those Chymists who speak plainest do by their Principles, certain primitive and Simple or perfectly unmingled bodies; which not being made of any other bodies, or of one another, are the Ingredients of which all those call'd perfectly mixt Bodies are immediately compounded, and into which they are ultimately resolved. (1661: 350)

Boyle's view of the elements seems, however, to have been *merely* sceptical, furnishing no positive framework for a detailed theory of the composition of substances (Siegfried 2002: Ch. 2). Boyle was sympathetic to atomism, which could clearly allow elements to be material components of substances, but he gave no account of how many kinds of atom there are, or how they, or their arrangements, would be linked to the identity of substances. As Ursula Klein puts it (1994: 170), Boyle's atoms are not *substance-specific* components.

During the eighteenth century there were a number of important changes in chemical theory and practice. Firstly there was a closer integration between compositional theory and expanding empirical knowledge (Siegfried 2002: Ch. 4). Secondly, the atmosphere came to be regarded as an active agent of chemical change (Siegfried 2002: Ch. 6). Most importantly, however, there emerged a conception of chemical substances as composed of stable substance-specific material components (Klein 1994; Klein and Lefèvre 2007: Ch. 2). This is most clearly expressed in the many affinity tables published in the eighteenth century. The first of these is Etienne-François Geoffroy's 'Table des différents rapports' (Geoffroy 1718). Geoffroy's table set out the orders of affinity between various substances: how one substance *X* can displace another *Y* from a compound *YZ* if it has a higher degree of affinity for the other component, *Z*. The reactions were assumed to be reversible, with *X*, *Y,* and *Z* treated as 'building blocks' (Klein 1994: 168–70) that persist as substance-specific material parts of substances throughout the changes he describes. To that extent they are '*relatively* stable entities' (1994: 170), in that their identity is conserved through specific classes of chemical change. Many such affinity tables were published in the eighteenth century, embodying the same general assumptions about composition (Klein and Lefèvre 2007: Chs. 8 and 9), assumptions which came to change the way chemists thought about hypothetical elements or principles like phlogiston (see Klein and Lefèvre 2007: 150–1). As a result, by the middle of the eighteenth century most chemists 'regarded the ultimate principles as a kind of physical component of mixts' (Klein and Lefèvre 2007: 114).

3.2. After Lavoisier

In the early nineteenth century, John Dalton proposed a form of atomism that was clearly relevant to compositional theory: to each of Lavoisier's elements there corresponds a particular kind of atom (to borrow Klein's phrase, Dalton's atoms were substance-specific, unlike Boyle's). They survive chemical change, underwriting the tacit assumption of the survival of the elements. Atoms of the same element are alike in their weight. On the assumption that atoms combine with the atoms of other elements in fixed ratios, Dalton's theory offered the sketch of an explanation of why, when elements combine, they do so with fixed proportions between their weights, and the fact that the same two elements may combine in different proportions to form distinct compounds.[6] Though Dalton's theory divided the chemical community, there is no doubt that in what Alan Rocke (1984) has called 'chemical atomism', it offered a highly influential conception of composition.[7]

The problem was how to estimate atomic weights. Although such tiny quantities could not be measured absolutely, they could be measured relative to a reference atom (the natural choice being hydrogen as 1), but how to set the ratio between the weights of different atoms? Dalton assumed that, if only one compound of two elements is known, it should be assumed that they combine in equal proportions. Water, for instance, was rendered as HO in the Berzelian formulae that chemists adopted over Dalton's own arcane notation. But Dalton's response to this problem seemed arbitrary (Rocke 1984: 35–40). Finding a more natural solution became pressing during the first half of the nineteenth century: more and more elements were being discovered, and the elemental composition of more and more chemical substances was being determined qualitatively. Disagreement over how to assign atomic weights could only add to the confusion.

Agreement on Cannizzaro's method for determining the atomic weights of the elements came after the Karlsruhe Congress of 1860 (Rocke 1984: Ch. 10). Dmitri Mendeleev attended the congress, and within ten years had published the first version of his periodic table (see Bensaude-Vincent 1986; Gordin 2004: Ch. 1). In doing so he made explicit the conception of the elements that motivated Lavoisier:

> [N]o matter how the properties of a simple body may change in the free state, *something* remains constant, and when the elements form compounds, this *something* has a material value and establishes the characteristics of the compounds which include the given element. In this respect, we know only one constant peculiar to an element, namely, the atomic weight. The size of the atomic weight, by the very essence of the matter, is a number which is not related to the state of division of the simple body but to the material part which is common to the simple body and all its compounds. The atomic weight belongs not to coal or the diamond, but to carbon. (Mendeleev 1869: 439)

The periodic table is a table not of simple substances, but of Paneth's basic substances, because Mendeleev correlated the atomic weights of elements with their properties across different states of chemical combination. Mendeleev, like Dalton, regarded the atoms of a particular element as alike in respect of their weight, but that assumption was discarded when isotopy was discovered in the early twentieth century.

That happened via the discovery of radioactivity by Henri Becquerel at the very end of the nineteenth century. It was not clear at first whether the underlying process was intrinsic to atoms: Mendeleev himself thought it to arise from a chemical interaction between heavy atoms and the ether (see Gordin 2004: Ch. 8). If it were a process of atomic disintegration, however, it made sense to analyse the decay products, new species whose place in the periodic table was unclear. Some of the decay products were found to be inseparable by chemical means from known elements, from which they had different atomic weights and radioactive properties (see Soddy 1966: 374–83). In 1910 Soddy proposed that the new elements should occupy the same place in the periodic table as the known elements they resembled, being not merely analogous, but chemically identical to them, and coined the term 'isotope' for the co-occupants. At first, isotopy seemed a relatively rare phenomenon, confined to a few species of heavy elements, but it proved to be more much more common following the development of mass spectrography (see Bruzzaniti and Robotti 1989), and (in 1913) H. G. Moseley's method for measuring nuclear charge, the property which isotopes share.

Many familiar elements turned out to have a number of different isotopes. In such cases the atomic weight as measured by earlier chemists reflected not a property of individual atoms, but the average of a population of atoms that was heterogeneous in respect of atomic weight. It could not be atomic weight that determined the chemical properties of an element, because its atoms may differ in their atomic weight, and atoms of chemically different elements may be alike in their atomic weight. Dalton and Mendeleev had turned out to be mistaken. There was some debate on how close were the chemical properties of different isotopes (see van der Vet 1979; Kragh 2000), but in 1923 the International Committee on Chemical Elements, appointed by the International Union of Pure and Applied Chemistry, enshrined the importance of nuclear charge as the determinant of the identity of the chemical elements (see Aston et al. 1923).

4. DISCOVERY?

We now have the resources to identify, in the terms of modern chemistry, the extensions of element names in the usage of past scientists. My central claim is that eighteenth-century chemistry and its nomenclature were shaped by interests in the qualitative patterns of particular kinds of chemical behaviour (combustion, calcination, acid-base reactions) and explaining them in terms of a particular conception of elemental composition. These patterns are determined by sameness and difference in nuclear charge, and quite insensitive to sameness and difference in atomic weight.

Consider Lavoisier's chemical neologisms 'oxygen' and 'hydrogen'. The gases to which he applied the names consisted mostly of oxygen and hydrogen atoms, respectively. Moreover, we have seen that he used the names, regardless of the state of chemical combination, to stand for components of substances that could survive chemical change. So, given the assumption that the same laws of nature applied then as now, his element names must have tracked a nuclear property that is invariant across chemical change, since non-nuclear properties of atoms such as their electronic configuration vary across different states of chemical combination. That leaves just three plausible candidates: nuclear charge (atomic number); atomic mass (weight, or mass number); or a combination of the two.[8] Atomic mass on its own can be rejected right away, since it bears no close relationship to chemical behaviour: isotopes of two different elements may share the same atomic weight.

Now it is true that Lavoisier's samples of oxygen and hydrogen would overwhelmingly have been alike in respect of both nuclear charge and atomic mass,[9] but it is nuclear charge rather than atomic weight that to a large degree determines an element's chemical behaviour, and that of its compounds. The chemical differences between isotopes are kinetic: they tend to undergo just the same reactions, but at different rates. The difference is a marginal effect: the heavier the isotopes, the less the isotopic difference (see Hendry 2006). Other substances known to Lavoisier as elements are much more diverse in respect of atomic weight: silver is approximately 52 per cent ^{107}Ag and 48 per cent ^{109}Ag; mercury a mix of around seven isotopes none of which forms more than a third of the total. Traces of other isotopes of oxygen and hydrogen would anyway have been present in Lavoisier's samples, since they occur naturally, and their presence would have been non-accidental given their chemical similarity to the dominant isotopes. It was the shared nuclear charge of these populations of atoms that caused them to be collected together as samples. Therefore in Lavoisier's usage, I conclude, element names apply to populations of atoms that are alike in respect of nuclear charge: 'oxygen' refers to whatever has a nuclear charge of 8, while 'hydrogen' refers to whatever has a nuclear charge of 1.

What, of broader philosophical significance, follows from this? One immediate consequence is that, if nuclear charge determined the extension of the element names even in Lavoisier's usage, and this fact was unknown until the early twentieth century, it must have constituted a discovery when it did become known.

Joseph LaPorte (2004: Ch. 4) has argued against precisely this discovery thesis, on the grounds that the reference of element names prior to the discovery of isotopy was indeterminate. His argument turns on a variant of Putnam's 'Twin Earth' thought experiment (2004: 103–8), in which, prior to the discovery of deuterium, Earth scientists travel to a planet where deuterium oxide (D_2O) fills the lakes and rivers, the locals calling it 'water.' But the Earth scientists notice a number of differences between this stuff and Earthian water, including the fact that Earthian fish brought along for the expedition die when placed in it. They conclude that Twin Earthian 'water' is not the same stuff as Earthian water, explaining its contrasting behaviour as arising from a different composition: it contains a new element (in fact deuterium), which is like familiar Earthly hydrogen in some ways, but different in others. Meanwhile, back on Earth, scientists discover deuterium, classifying it as an isotope of hydrogen, hence the same element. Since, argues LaPorte, the travelling scientists made no factual error, their conclusion must be one that might equally well have been made in the actual history of science. Prior usage was indeterminate over whether deuterium is hydrogen (and heavy water), so the actual IUPAC ruling in favour of nuclear charge rather than atomic weight must have been a decision rather than a discovery—one that changed the usage of 'hydrogen' along with the names of the other elements. LaPorte draws a wide-ranging conclusion from this argument: that Putnam's semantics for natural kind terms 'is, contrary to wide acclaim, useless in blocking instability. The causal theory leaves room for plenty of reference change' (2004: 118). This is because it allows that a natural kind term may be grounded in samples or examples of the kind 'by speakers whose conceptual development is not yet sophisticated enough to allow the speakers to coin a term in such a way as to preclude the possibility of open texture, or vague application not yet recognized' (2004: 118).

Some of LaPorte's conclusions can be resisted. The main argument does not withstand the conclusion that Lavoisier's and Mendeleev's usage of element names determinately picked out populations of atoms that were alike in respect of nuclear charge, but typically differed in respect of their atomic weight. Usage prior to the IUPAC decision was not indeterminate, and although (in the actual history of science) there was explicit discussion of the issue on two separate occasions after isotopy was discovered (see Kragh 2000), the decision to settle on nuclear charge was a decision only in a Pickwickian sense, to stick with the prior usage of element names. There were very good reasons to do so, since, from a chemical point of view, isotopes are so very alike. Moving on to LaPorte's wider conclusion, it seems too strong to say that Putnam's semantics is 'useless in blocking instability'. To be sure, it does not entail, all on its own, that there cannot be conceptual developments that bring with them referential change, but that is too much to ask of such a theory. In fact it would be odd if Putnam's theory ruled out referential instability in the absence of any particular assumptions about the historical usage of a term, or the intentions and interests of its users. Referential stability (and, for that matter, referential instability or indeterminacy) can be inferred only once the actual historical details of the introduction of a particular kind term are examined, along with the intellectual context in which it occurred. For this very reason philosophers of science tend to identify the value of Putnam's semantics to scientific realism as being that it allows the *possibility* of referential stability across theoretical or even conceptual change (see, for instance, Hacking 1983: 75; Psillos 1999: 280). Moreover, Putnam was aware that in order to fulfil its chief explanatory task, his theory presupposed a certain amount of conceptual and syntactical structure in the thought and language of users of a kind term (see, for instance, Putnam 1975: 269).

However, it must be conceded that LaPorte's arguments qualify the discovery thesis in important ways, and significantly complicate the connection between past and modern science. Our identification of the extension of element names in Lavoisier's usage was dependent on identifying the conceptual background to his thinking about elements. As we saw in section 3, Lavoisier drew on a conception of elements as material components of more complex substances whose general acceptance can be roughly dated to the mid-eighteenth century. So what can we say about the usage of element names on the part of earlier chemists? The issue doesn't arise for oxygen, hydrogen, or nitrogen, whose names were introduced by Lavoisier and his confederates, but what of lead, iron, copper, silver, gold, or sulphur which, along with many other elements, were known long before the eighteenth century? Since their discovery predated the core conception of the elements, the historical facts do not allow a simple application of the argument about oxygen. Furthermore, none of these elements was generally regarded *as* elements until later in the eighteenth century, so the core conception (had it been widely accepted earlier) would have been irrelevant to the extensions of their names. Speculation about what earlier chemists might have thought had they been provided with various pieces of information seems fruitless, which leaves us with LaPorte's conclusion that, prior to around 1700, it is indeterminate which populations of atoms were picked out by the names of these elements. The emergence of the core conception began to regiment the extension of element names only later.

Thus far I agree with LaPorte, but I would make two corrections. Firstly, his discussion of his 'Twin Earth' example presumes that conceptual refinement with respect to a kind term is a conscious decision occasioned by the empirical discovery of differences within its extension. As we have seen, however, the conceptual refinement that allowed chemists to refer determinately to populations of atoms of like nuclear charge long predated the discovery of nuclear charge and isotopy, and it was tacit. Determinate reference required not a conscious sharpening of usage in response to new discovery but the emergence of a particular conceptual apparatus for thinking about elements as material components of substances, and a project of understanding chemical change in compositional terms. I would also add that, subject to that emergence, it is a discovery that oxygen is the element with a nuclear charge of eight. Secondly, if one accepts my point that particular epistemic intentions regimented chemists' usage of element names, the significance of LaPorte's 'twin-earth' thought experiment is unclear. The scenario purports to show that usage of 'water' was indeterminate as to whether or not deuterium oxide counted as water until isotopy was discovered and the indeterminacy legislated away. But whose usage was supposed to be indeterminate? It seems highly implausible that, before the twentieth century, chemists' usage of 'water' was indeterminate as to isotopic extension, while the names of water's elemental components were regimented. Apart from heavy water, there is in any case no tradition in chemistry of distinct names for isotopic variants of compound substances. If the argument is meant to apply to vernacular usage, then it is unclear to me why it would have been changed in any way by the discovery of isotopy. Linguists have long claimed that 'water' in vernacular usage fails to track H_2O content, following surface appearances instead. The question of isotopic variance seems an impossibly distant refinement in this context.

This brings us to a further line of objection to the discovery thesis. Donnellan (1983: 98–104) describes another counterfactual history in which the 'Twin Earth' scientists work with just the same chemical substances as ours do, develop an atomic theory just as ours do, and come to recognize isotopy, just as ours do. Donnellan also imagines that some of the elements, as they naturally occur on Twin Earth, are dominated by a particular isotope,[10] and that the Twin Earthlings are 'more taken with' (1983: 100) atomic weight than with nuclear charge. Hence they do not regard as gold what we would regard as exotic isotopes *of gold* (e.g. ^{198}Au), and (he claims) there would have been no mistake in their doing so. The objection I want to concentrate on concerns the contingency of the classificatory interests of chemistry, and of the order in which particular scientific discoveries happened.[11] Some scientific results, so the objection runs, do not count as discoveries because they are contingent in important ways: they are overt decisions, or reflect the imposition of scientists' particular interests in more subtle ways.

What if Lavoisier and other scientists had had different interests, or a different background conception of composition? In fact it is not obvious that anything like modern chemistry would have emerged, because the standards of similarity and difference that shape a discipline influence what it discovers. None of this undermines the fact that nuclear charge is a real physical property that predated chemists' knowledge of it, or that fact that the patterns of behaviour it determines are a genuine feature of the causal structure of the world. There may be other patterns of behaviour in nature, and these may be of interest to other communities of scientists (actual or merely possible), but that is irrelevant to the extension of the element names as they were actually introduced.

Donnellan also considers the possibility that isotopes might have been discovered *before* the development of 'simpler atomic theory' (1983: 104) in which atoms were assumed to be alike in weight. In this case, the elements would have been isotopically pure, and there would have been many more of them:[12] to anyone adopting such a system, our elements would be mixtures. Now one must concede that the actual order of discovery had an effect on the conceptual development of chemistry, but it seems too strong to call it, as he does, a 'historical accident' (1983: 104). It is quite hard to think of counterfactual histories in which isotopy (difference *within* what we think of as elements) is discovered before the elements themselves are distinguished from each other, if we suppose the same laws of nature to be in force. In respect of chemical behaviour, the manifest respects in which different isotopes can be distinguished are subtle compared to the respects in which any particular element can be differentiated from others. Systematic investigation of isotopy required sophisticated separation methods whose development required a century's worth of chemical research after

Lavoisier, and also an integration of physical and traditionally chemical methods. Deuterium was discovered because there were small discrepancies (two parts in ten thousand) between the atomic weight of hydrogen as determined by chemical means, and through mass spectrometry. Early methods for isolating it exploited the differential rates at which $^{1}H_2O$ and $^{2}D_2O$ are electrolysed (see Farkas 1935: 115–7). Donnellan notes one 'big difference' (1983: 103) there can be between isotopes: some of them are radioactive while others are not. Despite the possibility of nuclear weapons, however, the effects of naturally occurring radioactivity are subtle because radioisotopes are so thinly spread. In the shape of Geiger counters and fast-breeder reactors, human artifice has, of course, magnified and concentrated the effects of radioactivity, but this ingenuity depends precisely on a detailed knowledge of chemical substances and how to isolate them in high degrees of purity. Even in concentrated form the effects are hardly manifest: uranium, all of whose isotopes are radioactive, was discovered in 1789, a full century before it occasioned the discovery of radioactivity. What if large enough quantities of pure samples of different isotopes had just been lying around, and chemists had been able to compare and contrast? In fact, given the similarities, they may well not have been able to distinguish them, but the situation is anyway fantastically unlikely, not only because some isotopes are rare or unstable, but also because the vast majority of natural processes just do not distinguish between isotopes.[13] This might seem like scientific pedantry, but the point is that sameness of nuclear charge is the similarity that brought together the particular populations of atoms that formed the actual samples studied by chemists, and in which the names of the elements were actually grounded. When isotopes become mixed, most natural processes have no tendency to separate them. Should we take seriously objections that centre on what scientists might have done in counterfactual situations which, though possible, are vanishingly improbable, given the nature of those processes?

5. NECESSITY AND ESSENCE?

Until now I have kept my discussion fairly close to chemistry and its actual history, and considered only the issues of semantics and classification. What prospects are there for an inference to the more substantial metaphysical thesis that (say) having the atomic number (or nuclear charge) 79 is what makes something gold? There is something of a critical consensus that the 'Kripke-Putnam' essentialist project is beset with difficulties. One major objection is that substantial metaphysical theses about kinds cannot be derived from mere semantic properties of their names: insofar as essentialism is derived from premises that apparently involve only semantic claims concerning rigidity and directness, substantial essentialist presuppositions must have been smuggled in (Salmon 1982: Chs. 5 and 6). This is one of many problems with the form of Kripke's argument from a supposed analogy between proper names and kind terms to the conclusion that 'gold is the element with atomic number 79' is a necessary identity known *a posteriori* (Kripke 1980; for recent criticism see Lowe 2008; Needham 2009a).

Another line of criticism questions how Kripke's and Putnam's thought experiments could hope to establish the necessity of 'water is H_2O' and 'gold is the element with atomic number 79'. Putnam's 'Twin Earth' example, for instance, strikes people in one of two ways: some conclude straight away that XYZ cannot be water because it is not H_2O, others just as surely that it *is* water because it is colourless, tasteless, boils at 100°C, and so on. What can one hope to establish from a thought experiment that elicits such divergent intuitions? Even if intuitions were more univocal, what can they establish anyway? Finally, even if one could establish the necessity of 'gold is the element with atomic number 79', there is the further difficulty of inferring that it marks an *essential* truth about gold. The problem is that there are too many metaphysical necessities (see Fine 1994). Consider a singleton set and its member: the set contains its member necessarily, but the member also belongs to the set necessarily. The necessities make neither set nor member prior, yet (argues Fine) the identity of the set depends on the identity of its member, rather than vice versa. There can be no question of deducing any form of essentialism from metaphysical necessity.

One response to all this is to eschew notions like necessity and essence when discussing scientific classification, but that seems unnecessary and undesirable to anyone who is sympathetic to scientific realism, and a fruitful interaction between metaphysics and the philosophy of science. Might one be more optimistic? Consider first the metaphysical necessity of 'gold has atomic number 79'. Rather than Twin Earth scenarios or semantic intuition, consider a kind of explanation that is typical in sciences that addresses complex situations: by considering the difference its absence would make, we identify one factor among many acting in concert as the cause of some part of an overall effect. To use an example I have developed elsewhere (see Hendry and Rowbottom 2009), chemistry textbooks describe the boiling point of water as 'unexpectedly high' (Gray 1994: 205). Why? The other hydrides in oxygen's group in the periodic table (hydrogen sulphide [H_2S], hydrogen selenide [H_2Se], and hydrogen telluride [H_2Te[) display a monotonic relationship between their boiling points and molecular masses (Gray 1994: 205),[14] and if this were extrapolated to water it would be a gas at room temperature. But water boils at 100°C because of hydrogen bonding, which 'modifies a great many physical and a few chemical properties' (Pimentel and McLellan 1960: 6). H_2O molecules are polar, and interactions between them are stronger than those between less polar molecules like H_2S, H_2Se, and H_2Te, which interact only via much weaker van der Waals forces. The explanation involves a contrast between water's actual behaviour and what would occur if hydrogen bonding were absent. Since hydrogen bonding occurs as a matter of physical law, the comparison involves actual and counternomic behaviour. Notice what is kept fixed in the counternomic situation: the overall number of electrons in the water molecule (10), and the nuclear charges: it would make no sense otherwise to determine its boiling point in the counternomic situation (or that of its counterpart) by extrapolating the trend from the other hydrides. If water's elemental composition is kept fixed in counternomic contexts, it is necessary in some strict way.[15] To that extent Kripke's and Putnam's intuitions seem right: what seems to exercise their critics is that they assume the epistemic perspective of scientific explanation. I leave aside how this necessity should be explained. Even if one is sceptical about how to evaluate claims about water in more exotic counterfactual or counternomic situations (e.g. whether it is possible for it to be a pink solid), here is an example with a clear epistemic purpose, and a situation that is not far from the actual. In this kind of explanation, it is very difficult to see what scientific sense could be made of allowing water's composition to vary, while other properties (e.g. its transparency or potability) are held fixed.

What of the connection between essence and (metaphysical) necessity? It is indeed a fallacy to *deduce* one from the other, but that doesn't mean that there is no connection. If it is metaphysically necessary that X has property P, this may well be part of an argument that it is essential to X that it has P, since the latter, if true, may explain the former. It depends on the case. In *this* case there are general grounds for taking nuclear charge to be the kind of property that determines the identity of a chemical substance because it is less like Fine's singleton set, and more like the member. Firstly, it is a causally important property that atoms carry with them from complex situation to complex situation in the actual world (the earlier example simply extends that to counternomic situations). Secondly, nuclear charge individuates the elements in another way: quantization of nuclear charge explains why the elements are countable (rather than a continuum).[16] Thirdly, in arguing that Lavoisier was able to refer to populations of atoms with a nuclear charge of eight, I first set out a threefold theoretical role for the elements in his thinking—the 'core conception'—and then identified nuclear charge as the property that realized that role. Not only do element names track nuclear charge in counternomic situations, but there is a good case for the priority of nuclear charge over chemical substances.

That brings us to the status of the core conception itself: is it *a priori* or *a posteriori*? If we take these terms to describe ways for something to be known, then it seems to me that it couldn't have been either in the eighteenth century, because neither Lavoisier nor anyone else *knew* that the core conception was the 'right' account of composition. There simply weren't any incisive criticisms of the alternative accounts of elemental composition. Perhaps the core conception, and alternatives like Aristotle's, should instead be seen as competing *a priori* possibilities for how composition might work. If it is known now that the core conception is 'right' (I think it is), then the reasons for that must involve the vindication of that conception of elemental composition by subsequent research in chemistry and physics. To that extent such knowledge is *a posteriori*, though it is not a *simple* empirical discovery because the subsequent scientific research proceeded in an intellectual environment that was already structured by the tacit (and unjustified) acceptance of just one of the possibilities for how composition might work. The project of accounting for the phenomena of chemistry through the core conception might have failed, just as chemists later came to understand that, in the case of the acids, the project of accounting for the behaviour of a class of compounds on the basis of their sharing a common component had failed (see Hendry 2005: 34). For that reason I would resist a deflationary account of the core conception, under which it is no more than the concept of element that happened to emerge during the eighteenth century, articulating a theoretical role for the elements that in the actual world turned out to be filled by nuclear charge (see, for instance, Jackson 2000: Ch. 3). Because the project might have failed, it must have embodied a substantial hypothesis.

NOTES

1. Indirectly, via his translators: Lavoisier himself, who wrote nothing in English, originally used the term 'principe oxigine' and, later, 'oxygène'.
2. The distinctness of Kripke's and Putnam's programmes has long been urged by Ian Hacking (1983: 82); for detailed accounts of the differences see Putnam 1990 and Hacking 2007.
3. Inclusion of the term 'uncontroversial' makes the application of this condition contextual: was there a live scientific debate in which the opposing sides differed over the interpretation of the experiment in question?
4. Caloric could be subject to quantitative measure even though it couldn't be weighed: Lavoisier used a calorimeter to gauge how much was released during the combustion of various substances, using this as a basis for estimating how much caloric they contain (1790: Ch. IX).
5. Following Paul Needham (2006, 2009b), from whom this brief sketch is derived, I use the term 'mixt' to indicate that Aristotle does not distinguish between compounds and homogeneous mixtures such as solutions.
6. Although see Needham 2004.
7. Just what chemical atomism was committed to is a subtle issue: Rocke makes clear that it should not be identified with ancient atomism or early modern corpuscularianism, though many anti-atomists based their criticisms on just such an identification.
8. Atomic number is just the ordinal numbering of a particular element in the periodic table. Nuclear charge is the physical property that, in 1913, was found to underlie that ordering (see §3). I will use them interchangeably.
9. ^{16}O and ^{1}H make up 99.76% and 99.99%, respectively, of naturally occurring oxygen and hydrogen.
10. This is in fact true of oxygen and hydrogen, as noted earlier. Naturally occurring Earthly gold is 100% ^{197}Au. Other known isotopes are unstable, with half-lives measured in days.
11. Donnellan also objects that it would be 'outrageously bizarre' (1983: 103) to suppose that 'psychological quirks' (1983: 103) like being 'more taken with'

atomic weight than with nuclear charge should determine the extension of kind terms in the usage of earlier scientists. True, but this is a consequence neither of Putnam's account, nor anything I have said here. It is the scientists in Donnellan's counterfactual history who have changed the extension of element names in line with their psychological quirks.

12. In fact Ida Noddack proposed just such a system of the elements in the 1930s. See Kragh 2000: 442–3.
13. This applies also to LaPorte's 'Twin Earth' scenario.
14. Molecular mass tracks the number of electrons, which influences the strength of van der Waals forces along with other factors like molecular shape. In a neutral molecule the number of electrons is of course equal to the total charge of the nuclei.
15. In fact I would go further: it seems very plausible that the water molecule's bond topology is also necessary.
16. Robin Le Poidevin (2005) bases upon this relationship an argument for the ontological reduction of chemistry. Le Poidevin is right to find the relationship metaphysically significant, although I do not see it as grounds for reduction (see Hendry and Needham 2007).

REFERENCES

Aston, F. W. et al. (1923) 'Report of the International Committee on Chemical Elements', *Journal of the American Chemical Society*, 45: 866–74.

Bensaude-Vincent, B. (1986) 'Mendeleev's periodic system of the elements', *British Journal for the History of Science*, 19: 3–17.

Boyd, R. (1990) 'Realism, conventionalism and "realism about" ', in G. Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam,* Cambridge: Cambridge University Press, 171–95.

Boyle, R. (1661) *The Sceptical Chymist*. London.

Bruzzaniti, G. and Robotti, N. (1989) 'The affirmation of the concept of isotopy and the birth of mass spectrography', *Archives Internationales d'Histoires des Sciences*, 39: 309–34.

Cassebaum, H. and Kauffman, G. (1976) 'The analytical concept of a chemical element in the work of Bergman and Scheele', *Annals of Science*, 33: 447–56.

Donnellan, K. S. (1983) 'Kripke and Putnam on natural kind terms', in C. Ginet and S. Shoemaker (eds), *Knowledge and Mind: Philosophical Essays*, New York: Oxford University Press, 84–104.

Dupré, J. (1993) *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press.

Farkas, A. (1935) *Orthohydrogen, Parahydrogen and Heavy Hydrogen*. Cambridge: Cambridge University Press.

Fine, K. (1994) 'Essence and modality', *Philosophical Perspectives*, 8: 1–16.

Geoffroy, E. (1718) 'Table of the different relations observed in chemistry between different substances—27 August 1718', translated in *Science in Context*, 9 (1996): 313–20.

Gordin, M. (2004) *A Well-Ordered Thing: Dmitrii Mendeleev and the Shadow of the Periodic Table*. New York: Basic Books.

Gray, H. B. (1994) *Chemical Bonds: An Introduction to Atomic and Molecular Structure*. Mill Valley, CA: University Science Books.

Hacking, I. (1983) *Representing and Intervening*. Cambridge: Cambridge University Press.

———. (2007) 'Putnam's theory of natural kinds and their names is not the same as Kripke's', *Principia*, 11: 1–24.

Hendry, R. F. (2005) 'Lavoisier and Mendeleev on the elements', *Foundations of Chemistry*, 7: 31–48.

———. (2006) 'Elements, compounds and other chemical kinds', *Philosophy of Science*, 73: 864–75.

———. (Forthcoming) *The Metaphysics of Chemistry*. Oxford: Oxford University Press.

Hendry, R. F. and Needham, P. (2007) 'Le Poidevin on the reduction of chemistry', *British Journal for the Philosophy of Science*, 58 (2007): 339–53.

Hendry, R. F. and Rowbottom, D. P. (2009) 'Dispositional essentialism and the necessity of laws', *Analysis*, 69: 668–77.

Jackson, F. (2000) *From Metaphysics to Ethics*. Oxford: Oxford University Press.

Klein, U. (1994) 'Origin of the concept of chemical compound', *Science in Context*, 7: 163–204.

Klein, U. and Lefèvre, W. (2007) *Materials in Eighteenth-Century Science: A Historical Ontology*. Cambridge, MA: MIT Press.

Kragh, H. (2000) 'Conceptual changes in chemistry: the notion of a chemical element, ca. 1900–1925', *Studies in History and Philosophy of Modern Physics*, 31B: 435–50.

Kripke, S. (1980) *Naming and Necessity*. Cambridge, MA: Harvard University Press.

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

Lavoisier, A. (1790) *The Elements of Chemistry*. Edinburgh: William Creech. Translation by Robert Kerr of *Traité Élémentaire de Chimie* (Paris, 1789).

Leicester, H. M. and Klickstein. H. S. (eds) (1952) *A Source Book in Chemistry 1400–1900*. Cambridge, MA: Harvard University Press.

Le Poidevin, R. (2005) 'Missing elements and missing premises: a combinatorial argument for the ontological reduction of chemistry', *British Journal for the Philosophy of Science*, 56: 117–34.

Lowe, E. J. (2008) 'Two notions of being: entity and essence', in R. Le Poidevin (ed.), *Being: Developments in Contemporary Metaphysics*, Cambridge: Cambridge University Press, 23–48.

Mendeleev, D. I. (1869) 'The relation between the properties and atomic weights of the elements', *Journal of the Russian Chemical Society*, 1: 60–77. Page references are to the translation in Leicester and Klickstein 1952, 439–42.

Needham, P. (2004) 'Has Daltonian atomism provided chemistry with any explanations?', *Philosophy of Science*, 71: 1038–47.

———. (2006) 'Aristotle's theory of chemical reaction and chemical substances', in D. Baird, L. McIntyre, and E. Scerri (eds), *Philosophy of Chemistry: Synthesis of a New Discipline,* Dordrecht, Netherlands: Springer, 43–67.

———. (2008) 'Is water a mixture? Bridging the distinction between physical and chemical properties', *Studies in History and Philosophy of Science*, 39: 66–77.

———. (2009a) 'Microessentialism: what is the argument?', *Noûs*, forthcoming.

———. (2009b) 'An Aristotelian theory of chemical substance', *Logical Analysis and History of Philosophy*, forthcoming.

Paneth, F. A. (1962) 'The epistemological status of the chemical concept of element (II)', *British Journal for the Philosophy of Science*, 13: 144–60.

Pimentel, G. and McClellan, A. (1960) *The Hydrogen Bond*. London: W. H. Freeman & Co.

Psillos, S. (1999) *Scientific Realism: How Science Tracks Truth*. London: Routledge.

Putnam, H. (1975) 'The meaning of "meaning" ', in *Mind, Language and Reality*, Cambridge: Cambridge University Press, 215–71.
———. (1990) 'Is water necessarily H_2O?', in J. Conant (ed.), *Realism with a Human Face*, Cambridge, MA: Harvard University Press, 54–79.
Rocke, A. J. (1984) *Chemical Atomism in the Nineteenth Century: From Dalton to Cannizzaro*. Columbus, OH: Ohio State University Press.
Salmon, N. (1982) *Reference and Essence*. Oxford: Blackwell.
Siegfried, R. (2002) *From Elements to Atoms: A History of Chemical Composition*. Philadelphia: American Philosophical Society.
Soddy, F. (1966) 'The origins of the conceptions of isotopes', in *Nobel Lectures, Chemistry 1901–1921*, Amsterdam: Elsevier, 371–99.
van der Vet, P. (1979) 'The debate between F. A. Paneth, G. von Hevesy and K. Fajans on the concept of chemical identity', *Janus*, 92: 285–303.

9 On the Abuse of the Necessary *A Posteriori*

Helen Beebee and Nigel Sabbarton-Leary

1. INTRODUCTION

Since Kripke famously argued that necessity and *a priority* can, in some circumstances, come apart, contemporary metaphysicians have increasingly appealed to the category of the necessary *a posteriori*. What they generally fail to do, however, is provide any *argument* for why the truths in question fall into this category. But arguments are needed. Even if we accept that Kripke's story holds for proper names and natural kind terms, it can by no means be taken for granted that the story extends to cover other cases. This paper rehearses the general argument that such arguments are indeed required, and discusses in detail one example of abuse of the necessary *a posteriori*: Brian Ellis's 'scientific essentialism' (SE), according to which the laws of nature are metaphysically necessary but knowable only *a posteriori*. Ellis grounds this alleged feature of laws in a conception of natural kinds that extends well beyond standard cases of molecular constitution and biological species ('water', 'gold', and 'tiger', to use Kripke's examples); we shall argue both that Ellis provides no convincing arguments for this extension, and that there are good reasons for thinking that no such arguments can be given.[1]

We shall proceed as follows. In §2, we draw out the consequences of Kripke's account of the semantics of natural kinds terms for anyone who wants to claim that the category of the necessary *a posteriori* is in fact much broader than Kripke suggests. In §3 we turn to Brian Ellis's scientific essentialism (SE), and argue that there are clear cases of Ellisian 'natural kinds' that fail to fit Kripke's model, and hence cannot be thought to generate the *a posteriori* necessities that Ellis claims to hold for natural kinds in general. In §4, we assess the consequences for SE, and argue that the lack of the relevant *a posteriori* necessities undermines Ellis's own account of *de re* necessity. Finally, in §5, we offer a brief diagnosis.

2. THEORETICAL IDENTITIES AND THE NECESSARY *A POSTERIORI*

Kripke argues that 'theoretical identity' statements, such as 'gold is the element with atomic number 79' and 'water is H_2O', are metaphysically necessary but knowable only *a posteriori*. As we pointed out in the introduction to this book, it is relatively uncontroversial that Kripke takes the terms on the left-hand sides of such 'identity' statements to be analogous to proper names such as 'Ehrich Weiss' and 'Harry Houdini' in that they are (i) non-descriptive, (ii) rigid, and (iii) introduced by ostensive baptism (or description, but where this description need have no bearing on the reference of the term). However, it is a matter of some controversy what Kripke thinks the semantic status of the terms appearing on the right-hand sides ('the element with atomic number 79'; 'H_2O') are. We favour an interpretation of Kripke that takes such terms to be a description of the essence of the natural kind in question, so that, for example, *having 79 protons in its nucleus* is the essence of gold, and *being constituted by molecules composed of two hydrogen and one oxygen atom* (perhaps along with some structural features of the molecules) is the essence of water. Thus the phrases that appear on the right-hand side of Kripke's paradigmatic theoretical identification sentences are (i) definite descriptions (used attributively), (ii) *de facto* rigid, and (iii) discovered and described by science rather than ostensively introduced.

The treatment of theoretical identifications as part of the category of the necessary *a posteriori* is thus—given the difference between expressions like 'the element with atomic number 79' and 'Harry Houdini'—not warranted by appeal to the similarities between proper names and natural kind terms alone. Insofar as Kripke offers any *argument* for the claim that facts like 'water is H_2O' are necessary yet knowable *a posteriori*, they are abbreviated analogues of Putnam's Twin Earth thought experiment (Putnam 1973): our initial identification of water was via its observable, characteristic properties, for example, its 'feel, appearance and perhaps taste' (Kripke 1980: 128), but as science improves we eventually discover that water is H_2O. Now Kripke invites us to imagine a scenario where there is a substance that has a 'completely different atomic structure from that of water, but resembled water in these [characteristic] respects' (ibid.); is this water? According to Kripke the answer is no. Just 'as there is a fool's gold there could be a fool's water' (ibid.). Given that water *is* H_2O, nothing lacking that atomic structure could be water.

The preceding interpretation of Kripke's account of natural kind terms is perhaps controversial. Nonetheless—and this is really all that matters for the purposes of this paper—it does at least show that Kripke provides the resources for two distinctively different routes for arguing that a given 'theoretical identity' statement is necessary *a posteriori*. One route is analogous to the route Kripke describes in the case of proper names. If one can plausibly maintain that the two terms in the 'identity' statement are *both* analogous to proper names (as perhaps 'gold' is, but 'the element with atomic number 79' clearly—we believe—is not), whose reference traces back to two distinct baptismal events which happen to name the same kind, then one has a good case for claiming that the relevant identity claim is necessary *a posteriori*. This route will play no role in what follows, since Ellis is explicitly concerned with essences.[2]

The other route is the route we attribute to Kripke earlier in the case of natural kind terms, and our starting point for the rest of this paper. Here, the term on the left ('gold') is analogous to a proper name and names a natural kind, and the term on the right ('the element with atomic number 79') specifies the essence of that kind. But one must *argue* that one's alleged case of a necessary *a posteriori* truth fits this model, and such an argument would seem to require two specific components:

Necessity: Consider the case of water. Let's grant that water has a straightforward underlying nature: actual samples of water are samples of a substance that is composed of molecules, each of which is composed in turn of two hydrogen atoms and one oxygen atom.[3] It does not follow, however, that 'water is H_2O' is *necessary*, because it does not follow that the underlying nature is water's *essence*. For of course it might be, for all that has been said so far, that the term 'water' is not analogous to a proper name: it might have a meaning such as 'whatever potable liquid is typically to be found in rivers and lakes and falls from the sky', in which case the claim that water is H_2O would be contingent rather than necessary. In order to rule this possibility out, we have to run a Twin Earth–style thought experiment: it is only the fact that we (allegedly) intuitively judge that XYZ is not water, despite meeting the above description, that justifies the claim that water's underlying nature is its essence, and hence that 'water is H_2O' is necessary.

A Posteriority: Even if we suppose that the relevant underlying nature *is* the essence of the kind in question, it still does not follow that we have a case of a necessary *a posteriori* truth. For it might be that the relevant rigid designator on the left of our theoretical 'identification' is introduced as a matter of stipulative definition, just as a bachelor is defined to be an unmarried man. Or it might be that it is what we shall call a 'descriptor':[4] a designator that has descriptive content that uniquely identifies the kind in question. Either way, the truth in question will be necessary but knowable *a priori*. Clearly this is not the case for 'water', since the word was around a long time before anyone knew the underlying nature of the substance the term refers to. But for many terms—and in particular for many of the terms typically assumed by essentialists to generate *a posteriori* necessities—it looks a lot more plausible that the term either is introduced as a matter of stipulative definition ('lepton' and 'Higgs boson', perhaps) or else is a descriptor ('ununseptium'). In order to rule this possibility out, then, one needs to provide an argument for the claim that the relevant term was not introduced by stipulative definition and is not a descriptor, so that it can plausibly be claimed that the essence of the kind is not simply the meaning of the kind term.

Kripke's own examples of natural kind terms, of course—'gold', 'tiger', 'water'—are all a part of ordinary language, and (as we just saw for 'water') it would—by Kripkean lights anyway—be implausible to maintain that their meaning is the underlying essence of the kinds in question. Hence his argument focuses on Twin Earth–style thought experiments, and he does not bother to argue explicitly for *A Posteriority* in the sense just described. The situation is rather different for many contemporary essentialists, who are largely concerned with the *fundamental* joints in nature, which are likely to be found in the classifications of physics rather than those of ordinary language users. Given these philosophers' conception of natural kinds, hardly any natural kinds have ordinary-language names like 'gold' attached to them. Hence, as we shall see in the case of Ellis, the argumentative lacunae in their views tend to be found in the absence of any defence of *A Posteriority* rather than *Necessity*.

However, since an argument for *A Posteriority* involves arguing that a given term is not introduced by stipulative definition and is not a descriptor, such an argument will always open up the possibility described in *Necessity*: that what is discovered *a posteriori* is merely the underlying nature of the kind in question, rather than its essence; and this possibility will in turn need to be removed using a Twin Earth–style thought experiment. So a complete argument for necessary *a posteriority* will, in the end, need to cover both bases.

3. ELLIS ON NATURAL KINDS AND LAWS

Brian Ellis's *Scientific Essentialism* (2001) presents a full-blown essentialist theory of the fundamental nature of reality. For Ellis, there are several kinds of natural kind: substantive kinds (proton, hydrogen atom), dynamic kinds or kinds of processes (β-decay, refraction, electromagnetic radiation), and property kinds (having a rest mass of two grammes, having spin ½). Natural kinds form a hierarchy, so that, for example, *nitrogen* is one 'infimic species' (and hence itself a natural kind) of the more general kind, *element*, and hydrogen atoms are members of one infimic species of the more general kind *atom*.

Natural kinds, for Ellis, ground laws of nature, and these laws are necessary *a posteriori*. Thus, for example:

> The laws of electromagnetism . . . must hold of electromagnetism in any world in which electromagnetic radiation may exist. The laws had to be discovered empirically, of course, so they are *a posteriori*, in the way that all empirical generalizations are. But what has been discovered is the essential nature of such radiation—that is, the properties and structure that any radiation must have if it is to be electromagnetic radiation. The laws of electromagnetism are thus necessary *de re*. (2001: 226)

Similarly for substantive and property kinds:

> . . . it is a necessary truth that a thing of kind *K* has the property *P* if *P* is an essential property of *K*. It is, of course, *a posteriori* what properties are essential to a given kind. Therefore, the proposition that things of the kind *K* have the property *P* is what I call 'really necessary'. If *P* is a natural dispositional property, then it is also a necessary truth that anything having the property *P* must be disposed to behave in certain ways in certain circumstances just in virtue of having this property. Of course, we have to discover empirically what kinds of dispositional properties exist. But if anything has the property *P*, it must be disposed to behave in a *P*-wise fashion, just in virtue of being a thing of this kind. Therefore, if the laws of nature are propositions stating facts of this sort, then they too are really necessary. (2001: 219)

The first point that needs to be made about Ellis's position is that he simply takes it for granted that it is '*a posteriori* what properties are essential to a given kind': the essential natures of natural kinds are for scientists to discover by empirical investigation. This is a very large assumption indeed. Granted that Kripke shows this to be so for the kinds *gold*, *water,* and *tiger*—and perhaps we can generalize to cover those basic chemical kinds and biological species for which we have names in ordinary language ('diamond', 'dog', 'charcoal', and so on)—we cannot massively expand the remit of natural kinds terms and assume without argument that what goes for Kripke's natural kinds goes for all natural kinds, given our expanded conception of them. After all, while Kripke is undoubtedly giving an account of natural kind terms, in effect for Kripke what *makes* something a natural kind term is the fact that it obeys the requirements of his theory. If one wishes to expand the extension of 'natural kind' to cover cases that are manifestly highly dissimilar from Kripke's examples (so that it covers, for example, leptons and refraction), one needs to *show* that the terms used to denote these kinds obey Kripkean semantics, if one wishes to preserve Kripke's claim about *a posteriority.* This is not something that Ellis does. Indeed, he explicitly says that his 'concerns are different from Kripke's. Kripke's essentialism was developed in relation to theories of reference and identity. Scientific essentialism [that is, Ellis's view] derives from an examination of the scientific practice of theoretical identification' (2001: 54). But of course if Ellis means something like what Kripke means by 'theoretical identification'—which he does, because he holds that that the relevant identifications will be necessary *a posteriori*—then Ellis should be just as concerned as Kripke is with reference and identity: without an argument that what goes for gold and water goes equally for leptons and refraction, his claim that natural kinds (as he understands them) *in general* generate *a posteriori* necessities is unwarranted.

Of course, the claim may yet be plausible or warrantable, even though Ellis himself does not make the argument. But consideration of some particular cases of chemical and process kinds will demonstrate that there are at least some natural kinds (in Ellis's sense) that manifestly do not generate *a posteriori* necessities. We shall focus primarily, in §3.1, on the case of natural kinds of substance, and deal only briefly, in §3.2, with process kinds.

3.1 Substance Kinds

Recently (in June 2009) the International Union of Pure and Applied Chemistry (IUPAC) confirmed the discovery (or rather, manufacture) of element 112 by Sigurd Hoffman and his team at the Centre for Heavy Ion Research in Germany, which is soon to be added to the periodic table. Chemists concerned with nomenclature—the systematic naming of substances—have introduced a decisive system for the introduction of 'temporary designators',[5] used to name elements for which there is evidence they exist, but where that evidence falls short of conclusive proof. In such cases the IUPAC advise that an element name be 'derived directly from the atomic number of the element' (Connelly et al. 2005: 47). Element 112 was first reported in the mid-90s, but IUPAC standards deem that a single reported discovery is insufficient to confirm the existence of a new element. Hence a temporary designator was introduced to refer to element 112. The systematic rules are based upon ten numerical roots. Each numeral of the atomic number of an element is replaced with the corresponding letters. Thus '1' is replace by 'un' for both the first and second numerals, and '2' is replaced by 'bi', and the series of letters is 'terminated by "ium" to spell out the name' (ibid.). Finally, following convention, the 'i' of 'bi' is elided to give us the element name 'ununbium'.[6]

Similar examples using descriptors are also available for more complex kinds. Consider the compound consisting of molecules of PCl_3. According to IUPAC there are three distinct systems of nomenclature that can be used to generate three distinct names for the same compound, and our choice of which to use is determined, in part, by how much information we are intending to convey with that name. The simplest *compositional nomenclature* employs a system that recommends names 'which are based solely on the composition of the substance' (Connelly et al. 2005: 5) and stipulates certain grammatical rules 'to specify the ordering of components, the use of multiplicative prefixes, and the proper ending for the names of electronegative components' (ibid.: 6). On this system PCl_3 comes out as phosphorus trichloride. The second, more complex, system is *substitutive nomenclature*, and is 'based on the concept of a parent hydride modified by substitution of hydrogen atoms by atoms and/or groups' (ibid.). The nomenclature specifies rules for naming the parent compound and substituent atoms and/or groups of atoms. In this case PCl_3 comes out at trichlorophosphane. Finally, the most complex system is *additive nomenclature*, where compounds are treated as the 'combination of a central atom or central atoms with associated ligands' (ibid.: 7). 'Ligand' is the term used to denote any substance, be it an atom or a molecule, that is bonded to the central atom. The grammatical rules of the *additive nomenclature* 'provide ligand names and guidelines for the order of citation of ligand names and central atoms names, designation of charge or unpaired electrons . . . [and] designation of spatial relations' (ibid.). On this system PCl_3 comes out as trichloridophosphorus.

Each of the names—'phosphorus trichloride', 'trichlorophosphane', and 'trichloridophosphorus'—communicates some basic information even to laymen who are not familiar with the specifics of the grammar of each individual system. Take 'trichloridophosphorus', for instance. As complex as this name is, it only takes a passing acquaintance with the periodic table to know that the name refers to a compound consisting of three parts chlorine and one part phosphorus. More important, perhaps, is the function these names perform for those individuals who *are* competent with the grammar. Names produced using the additive system will allow someone *au fait* with the grammar to construct a representation of the molecule, including its charge and the spatial relations between the constituent atoms.

What both of these examples illustrate is that some—and indeed clearly most—chemical names are *not* introduced using a Kripke-style name-acquiring transaction. Rather, they are generated using a complex set of rules and grammar, and clearly encode descriptive information. In other words, they are descriptors. As a result, a theoretical identity sentence such as 'ununbium is the element with atomic number 112' and 'trichloridophosphorus is PCl_3' is something a chemist can come to know *a priori*. We simply could not have discovered that trichloridophosphorus was *not* PCl_3, given the way that the name was introduced, nor that ununbium was *not* the element with atomic number 112. Many natural kind terms, then, do not adhere to the orthodox Kripkean model.

Chemical terms derived from IUPAC rules are not the only names of putative natural kinds that fail to fall within the remit of the Kripkean story; there are plenty of other cases that, at the very least, cannot simply be assumed without argument to generate *a posteriori* necessities. Consider the Higgs boson. Clearly there was no initial baptism, akin to the naming of water, for the Higgs boson. Rather, the existence of the Higgs boson is a hypothesis designed to explain why the photon has no mass while the W and Z particles (responsible for weak nuclear force) have huge masses—a hypothesis that has yet to be confirmed, thanks to teething problems with the Large Hadron Collider at CERN. So, for example, the statement 'the Higgs boson (if it exists) is responsible for the masses of W and Z particles' would seem to be knowable *a priori*: we know *a priori* that any particle that is discovered that fails to account for the masses of W and Z particles will not be the Higgs boson but something else. The same goes for 'the Higgs boson has no spin': since a boson is *defined* as having no spin, we know *a priori* that any particle that is discovered that has spin will not be the Higgs boson.

Of course, there are plenty of things that physicists might find out *a posteriori* about the Higgs boson, assuming that they discover that it actually exists; and one might argue that such truths *will* be good candidates for members of the category of necessary *a posteriori* truths about the essence of the Higgs boson. For example, the Standard Model of particle physics does not predict the mass of the Higgs boson. So, say, 'the Higgs boson has mass 120 GeV (gigaelectronvolts)' is not knowable *a priori*. Unfortunately, however, it is not obviously *necessary*: we would need an *argument* in order to establish that the claim 'the Higgs boson has mass 120 GeV' is necessary if true. As we've seen, what would be needed in order to establish this would be a Twin Earth–style thought experiment. But we cannot simply *assume* that the Higgs boson's mass is part of its essence, in the kind of full-blooded way that generates *a posteriori* necessity.

Thus at least some of the kinds identified by Ellis as 'natural' kinds turn out to have essences that are defined rather than discovered: our candidate 'theoretical identifications' are knowable *a priori*, and we have no grounds for thinking that those underlying features that are plausibly taken to be discovered *a posteriori* (for example, the mass of the Higgs boson) are a part of the kind's essence.

It might be objected that the situation is not as clear-cut as we have suggested. The claim that 'ununbium is the element with atomic number 112' is knowable *a priori* entails that it could not be discovered to be false; however, this is disputable. Imagine, for example, that the element we were calling 'ununbium' turned out, long after the term had come to be widely used (and perhaps part of ordinary language because [what we had been calling] ununbium turned out to have properties that are important outside the chemistry lab), to have 113 protons rather than 112 protons in its nucleus. Surely in this case we would want to say that ununbium turned out not to have atomic number 112, in which case it cannot be *a priori* that ununbium is the element with atomic number 112, since we can imagine the theoretical identification turning out to be false. Hence, contrary to what we have argued, 'ununbium is the element with atomic number 112' is in fact necessary *a posteriori*.[7]

Such a claim would, of course, depart from Kripke's original metasemantic story, since in the aforementioned case there is no name-acquiring transaction of the Kripkean variety. Rather, the thought would be that, whatever the genealogy of the name, at some later time it *comes* to, as it were, lose its descriptive content and become a name that directly refers to the element in question.

Our response is to accept the thought experiment, but deny that it shows that 'ununbium is the element with atomic number 112', *as the term 'ununbium' is currently used*, is known only *a posteriori*. Clearly if it was discovered, right now, that Prof. Hofmann and his team had actually manufactured samples of the element with atomic number 113, IUPAC would determine that what they had actually manufactured was ununtrium, and not ununbium at all, and that ununbium does not yet exist (or, if it occurs naturally, has never been discovered). Of course, this means that previously uttered claims involving 'ununbium' would all turn out to be false. Or, if that seems implausible, one might say some such claims, in some contexts, involved a referential rather than attributive use of the term, so that 'Great, we've manufactured ununbium!' would be false, but, say, 'Here's that sample of ununbium you asked for' would be true. This would not undermine the claim to *a priority*, since the same holds for terms like 'bachelor': it is of course knowable *a priori* that all bachelors are unmarried, but one can, in certain circumstances (if Donnellan [2008: 268] is right, at any rate), use the term 'bachelor' referentially, as in the utterance of 'the bachelor in the corner is wearing a terrible suit', where it is clear which man the speaker intends to refer to, but they are mistaken about the man's marital status.

In other words, we can accept that the meaning of 'ununbium' (or for that matter, 'Higgs boson', and perhaps even 'phosphorus trichloride') could, in principle, change so that the term comes to be directly referential. But that does not undermine the claim that the term *actually* has descriptive content that renders the relevant theoretical identity statement knowable *a priori*.

3.2 Process kinds

Ellis holds that what (allegedly) goes for natural kinds of substance goes also for natural kinds of process. He says:

> A natural kind of process that is a display of a dispositional property has a certain real definition. And it is one of the primary objects of science to try to discover what the real definitions of the various natural kinds of processes are. In the case of any simple causal process, the real essence will be a dispositional property, and the scientific problem will be to specify precisely what this dispositional property is. In general, the real essence of a causal process of a given natural kind will be specifiable counter-factually by the kind (or kinds) of circumstance *C* in which it would be triggered, and the kind (or kinds) of outcome *E* which would . . . result, if there were no interfering or distorting influences. In the simplest kind of case . . . the dispositional property may be uniquely characterized by an ordered pair <*C, E*> where '*C*' denotes a kind of circumstance and '*E*' a kind of event. If *x* is an object that has this dispositional property, then *x* may be said to have the power, capacity, or propensity to *E* in circumstances *C*. However, it is not an *a priori* matter what the real essences of the natural kinds of processes are, and what is being determined is not the meaning of a dispositional term. (Ellis 2001: 124)

It is clear, then, that Ellis thinks that there are necessary *a posteriori* truths to be had in the domain of dispositional properties and the causal processes to which they give rise. A dispositional property *P* has an essence, *E* (so that, in the simplest case, for an object to have *P* is for it to be such that outcome *O* would occur, were the object to be placed in circumstances *C*), where it is an *a posteriori* matter, to be discovered by scientific investigation, that the essence of *P* is *E*, but where it is metaphysically necessary that *P* is the dispositional property with essence *E*. We shall argue, briefly, that the claim that the category of the necessary *a posteriori* can be extended to dispositional properties, and thereby to causal processes, in this way is highly implausible.

Recall our lessons from Kripke. We get theoretical identifications—the relevant class of necessary *a posteriori* truths—just when we have a general term (such as 'gold') on the left-hand side of the 'identity' statement that has no descriptive content or stipulative definition, and a specification of the essence of the kind thus named on the right-hand side.

Note first that in the passage just quoted, Ellis advances the claim that dispositions have 'real definitions' or 'essences' that are not knowable *a priori* without the slightest hint that a Kripkean story about the naming of dispositions is required in order to justify this claim. In his 'simplest case', our dispositional property 'may be uniquely characterized by an ordered pair <*C, E*>, and if object *x* has this property, then it 'may be said to have the power, capacity, or propensity to *E* in circumstances *C*'. This would appear to deliver the metaphysically necessary truth: 'an object with property <*C, E*> has the power to *E* in circumstances *C*', or perhaps 'the property <*C, E*> *is* the power to *E* in circumstances *C*'. Metaphysically necessary this may be, but knowable only *a posteriori* it most certainly is not: it is a matter of orthographical stipulation that property <*C, E*> *is* the power to *E* in circumstances *C*, and so this is uncontroversially knowable *a priori*. (Of course, whether such a property *exists* is not knowable *a priori*, but this is not what Ellis is claiming. He is claiming that the *essence* of the property is not knowable *a priori*.)

Many properties do, of course, have names that are not merely orthographically distinct ways of describing their essence. Take solubility, for example—a property that Ellis takes to be 'a real disposition, for the process of solution is a natural kind of process' (2001: 125). Let's say that the definition of solubility (in some substance *S*) is the power to form a homogeneous mixture with *S* with particle sizes at the molecular or ionic level. Grant, then, since this is our toy definition, that it is metaphysically necessary that this is so. The question is, is it knowable only *a posteriori*, or have we merely stipulatively defined what it is to be soluble, so that our metaphysically necessary truth is knowable *a priori*?

It is not obvious, straight off the bat, what the answer to this question is. One might argue, in defence of Ellis, that a Kripkean story can be told about solubility. For example, one can imagine ordinary language users or proto-scientists noticing the (normally) observable difference between, say, what happens when one mixes salt with water (producing a solution) and what happens when one mixes sand with water (producing a suspension): in the second case but not the first, the resulting mixture is murky and its constituent substances separate out if you leave the mixture to stand. And one can imagine our proto-scientists dubbing the disposition to produce a mixture of the first kind 'Disposition *S*'. So we have a potentially Kripkean story here: an initial baptism based on observable features, where the underlying 'essence' of the disposition is still waiting to be investigated by science (since one can—fallibly—successfully identify the effects of possession of Disposition *S* in the presence of its manifestation conditions without any grasp or knowledge of the fact that the essential difference between a solution and a suspension is the size of the particles that constitute the mixture). So there is a prima facie case for thinking that Disposition *S* is solubility, and that there are *a posteriori* necessary truths to be had about its underlying nature (perhaps that Disposition *S*—that is, solubility—results in molecule- or ion-sized particles distributed homogeneously in a mixture).

So far, so good. Unfortunately, however, matters are rather more complex than is suggested by the toy story just given. Solutions are often defined in terms of the size of the particles in the resulting mixture: up to 2 nanometers for a solution, and over 1000nm for a suspension. In between the two, there is the category of *colloid*: a mixture where the particle size is between 2 and 1000nm. Many colloids (unlike solutions) can be separated into their constituent substances by filtration (dissolved particles are too small for this), but unlike suspensions they do not separate out naturally, because the colloidal particles (like dissolved particles) are small enough for Brownian motion. Blood and mayonnaise, for example, are colloids. Also, solutions, colloids, and suspensions need not be mixtures of solids and liquids, or liquids and liquids: carbonated water, water vapour in the air, and brass (an alloy of copper and zinc) are all solutions; pumice and smoke are both colloids. Now, which disposition did our initial baptism of 'Disposition *S*' denote, exactly? It seems highly unlikely that this question has a determinate answer—and if it does, it seems highly unlikely that the answer is 'solubility'. First, the observable features on the basis of which the initial baptism was made (the fact that salt, unlike sand, forms a transparent liquid that does not separate out on standing) are neither necessary nor sufficient conditions for something's being soluble. Of course, this is true of natural kind terms generally on the Kripkean story—it is neither necessary nor sufficient for something's being a sample of water that it is potable, found in a river or a lake, and so on. But such features are at least reasonably typical (as opposed to necessary) and distinctive (as opposed to sufficient) of water. In the case of solubility, the relevant features are neither typical nor distinctive: water vapour in air is not a transparent liquid, and nor is brass, but they are both solutions, and colloids as well as solutions fail to separate on standing. It is therefore hard to see how to make the case that the disposition denoted (if any) was solubility. This would be a little like claiming that someone who had only ever been exposed to one kind of mammal—cats, say—and who baptised that kind of entity (pointing to a cat) '*M*' established that '*M*' denotes the kind *mammal*. If '*M*' refers to anything, then its most plausible reference is the kind *cat*.

Perhaps, by analogy with the cat case, one could argue that a more restricted disposition, in the right ballpark, was denoted—solubility in water, say. But solubility in water is not solubility. So if 'solubility' *simpliciter* is to count as a kind term capable of generating *a posteriori* necessary truths, this will not help.

Second, we have so far glossed over a difficulty for the very possibility of 'baptising' a disposition in the first place. One cannot, it seems to us, directly baptise a disposition, because dispositions—when they are not being manifested—have no observable features whatsoever (analogous to be a colourless, potable liquid, say), which one can use to fix the reference. I cannot point to a sample of salt and say 'let the dispositional property had by that sample be named "solubility"', for of course it will then be completely indeterminate which dispositional property I am attempting to pick out. So it seems that the only way to baptise a disposition would be indirectly, via denoting the process, or perhaps the product of the process, that the disposition gives rise to. Indeed, this is how we told our preceding toy story about solubility: we start by picking out a process (dissolving), or perhaps the result of that process (a solution), and then denote the disposition by calling it the disposition to do something of *that* kind, or to produce something of the same kind as *that*. But this would seem to require not only that solubility is a natural kind, but also that both the process of dissolving and the result of that process—a solution—is also a natural kind. The latter mechanism for baptising solubility (or solubility in water) obviously requires this: if solubility is the disposition to produce something of *that* kind, then 'that' must directly and non-descriptively denote a kind—that is, it must denote a Kripkean natural kind. The former mechanism—where the process rather than the product is directly denoted—would also seem to require that *solution* is a natural kind; at any rate, it would be strange if both solubility (the disposition) and dissolution (the process) were natural kinds but the product of the process (solution) were not.

This creates a serious problem for Ellis's account, because it is in serious tension with his 'hierarchy requirement': 'if anything belongs to two different natural kinds, these natural kinds must both be species of some common genus' (2001, 20). Consider the term 'water'. Most likely, all actual samples of H_2O have other substances dissolved in them (most obviously, chlorine in tap water and salt in seawater). So are these in fact (impure) samples of H_2O, or (pure) samples of solutions? The hierarchy requirement rules out the possibility of answering 'both' to this question, since of course there is no common genus of which H_2O and *solution* are both species. So the answer must be one or the other. But (most, and probably all) samples of (what we call) water just *are*, uncontroversially, solutions. But then, by the hierarchy requirement, they cannot also be samples of H_2O. And of course in that case it's going to turn out that 'water' does not in fact refer to the natural kind we thought it referred to: it refers to the natural kind whose essence is that it is a solution, and not the natural kind whose essence is that it is H_2O!

We conclude that the example of solubility—one of Ellis's examples of a 'real disposition' (2001: 125)—is not a kind of disposition that generates any interesting *a posteriori* necessities, because the term 'solubility' simply does not fit the required Kripkean story. Of course, it might be that Ellis should not have said that solubility in particular is a real disposition; he might be wrong about solubility but right that there *are* real dispositions and natural kinds of process about which there are metaphysically necessary truths that are knowable only *a posteriori*. Our point, as with substance kinds, is that this needs to be shown rather than assumed.

Before considering the consequences of the fact that some theoretical identifications are knowable *a priori* for Ellis's overall position, it is worth noting that it is arguably not merely *some* Ellisian natural kind terms that fail to generate the required *a posteriori* necessities, but the vast majority. Kripke's own examples—'water', 'tiger', 'gold'—all designate objects or substances that are easily observable and (fallibly) individuated by ordinary people, and this is what makes the claim that such terms are directly referential plausible. Very, very few Ellisian natural kinds—the allegedly natural joints in nature that scientists aim to uncover—fall into this category. Scientific investigation—particularly in physics and chemistry, where (given that Ellis holds that biological species are not natural kinds) most of the Ellisian kinds are to be found—rarely involves finding some unidentified object or substance, giving it a name, and *then* investigating its nature. This point also applies to alleged natural kinds of process and to the dispositions that allegedly give rise to them: 'solubility' is not at all like 'gold', and nor are 'refraction' or 'electromagnetic radiation'. So our claim is not that kind terms such as 'ununbium', 'Higgs' boson', and 'solubility' are counterexamples to a rule that holds in all but a few cases; rather that it is ordinary-language terms such as 'water' and 'gold' that are the exceptional cases.

4. THE CONSEQUENCES FOR SCIENTIFIC ESSENTIALISM

How much trouble does this make for Ellis's overall essentialist picture? Quite a lot, we shall argue, because the philosophical core of Ellis's position is the view that 'natural necessities' are 'grounded in the world' (2001: 248). His position thus contrasts with the Lewisian conception of modality, according to which necessity is an inter-world, rather than intra-world, phenomenon: the necessary status of a truth is grounded not in the nature of the actual world—in which, in and of itself, there is no necessity to be found—but in the relationship between the actual world and other possible worlds. So, for example, for Lewis the necessity of the laws of nature ('physical' necessity rather than metaphysical necessity) is secured by fiat: a proposition is physically necessary, by definition, just in case it holds at all possible worlds with the same laws of nature as the actual world. Ellis's position also contrasts with David Armstrong's, according to which, while natural necessity is a real, fundamental feature of the actual world, its existence—and which universals it relates—is contingent (see Armstrong 1983). So natural necessity is, as it were, separable from the intrinsic natures of things: it is a kind of glue that binds things together.

Ellis, by contrast, says:

> To ground natural necessities in the world, it is necessary to develop an ontology of things capable of sustaining causal and other modal relationships. To do this, it is necessary to ground the laws of nature somehow in the world. To deal with the bulk of laws of nature—the causal and statistical laws that describe the powers and structures of things belonging [to] natural kinds—it is sufficient to recognize that these things have their kind essences essentially. (2001: 248)

For Ellis, then, natural necessity is grounded in essences: where the 'powers and structures' of things are essential to the kinds that those things are members of, those essences will sustain 'causal and other modal relationships'. Natural necessity is thus, for Ellis, full-blooded, *de re* metaphysical necessity, or 'real necessity' as he sometimes puts it. Natural necessity arises out of the intrinsic, essential natures of the inhabitants of the actual world, and not from any relationship between the actual world and other possible worlds (as Lewis has it); nor is it (as Armstrong has it) an additional item of ontology—a kind of metaphysical glue.

The problem for Ellis arises when we ask what the argument is for the claim that there *are* essences—of substantive, property, and process kinds—of the sort that will generate 'real' necessities. Ellis says:

> Analytic propositions are true in virtue of the meanings of words—that is, they depend for their truth on some conventionally established criterion for including something in some linguistically defined class. Metaphysically necessary propositions, on the other hand, are true in virtue of the essential natures of things—for example, they state correctly, or otherwise depend for their truth on, what makes something a thing of the natural kind it is. (2001: 235)

How are we to understand this claim in the context of the examples of chemical kind terms just described? 'Ununbium is the element with atomic number 112' and 'phosphorus trichloride is PCl_3' are, we argued earlier, analytic: all that is needed in order to know the truth of these claims is some rudimentary knowledge of the mechanics of chemical nomenclature. But, given Ellis's characterization of the difference between analytic and metaphysical necessity, it would seem that he is committed to saying that ununbium and phosphorus trichloride are not natural kinds. Since the relevant truths depend merely on 'some conventionally established criterion for including something in some linguistically defined class', they are *not* 'true in virtue of the essential natures of things' and thus (since all natural kinds have essences) are not truths about (the essences of) natural kinds. Gold and water, by contrast, *would* seem to have essences and thus be natural kinds, since 'gold is the element with atomic number 79' and 'water is H_2O' are not analytic.

This would be a curious position, to say the least, for Ellis to endorse: after all, gold and ununbium are both elements, and water and phosphorus trichloride are both compounds. So if gold is a natural kind with an essence that generates *de re* necessity, surely ununbium is too; similarly for water and phosphorus trichloride. And it is very clear that Ellis would want to endorse this claim. But the problem is that he cannot, apparently, endorse it, because he is committed to the view that analytic truths cannot be truths about essences.

In fact, Ellis goes on to offer a criterion for distinguishing between analytic and metaphysical necessity that, on the face of it, sidesteps this problem:

> . . . one technique is to abstract from the descriptive language used to refer to [some class of objects], and replace the general name used with an ostensive 'kind-referring' expression, such as 'stuff of this kind' or 'things of this kind'. If the necessity survives this process, then we know that it cannot be grounded in the descriptive language we had been using. 'Water is H_2O', for example, clearly survives this test, because 'stuff of this kind is H_2O', said pointing to a glass of water, is no less necessary than 'water is H_2O'. If there is any doubt about it, then it can only be a doubt about what the intended object of reference is [e.g. about whether one is referring to the glass or its contents], or ignorance about what its nature is. (2001: 235–6)

Ellis's technique appears to solve the problem we just raised, since, arguably, what goes for water goes equally for phosphorus trichloride and ununbium: arguably 'stuff of this kind', said pointing to a sample of PCl_3, is no less necessary than 'phosphorus trichloride is PCl_3'.

So far so good. But now a new problem emerges: Ellis's technique for marking out the realm of real necessity is clearly far too permissive. The contrast he draws is with replacing 'bachelor' with 'a person of this kind', the idea being that the latter expression fails to fix reference to a particular class. And this is because 'bachelors are not bachelors in virtue of their intrinsic properties or constitutions': 'Examine any given bachelor as thoroughly as your please; you will never discover the intended reference of the word "bachelor" as a result of such an investigation' (2001: 236).

Unfortunately, however, Ellis here sets up a false dichotomy. It is, of course, true that bachelors are not bachelors in virtue of their intrinsic properties or constitutions. But plenty of members of other intuitively non-natural kinds *are* members of those kinds in virtue of their intrinsic properties or constitutions. Imagine, for example, that you point to a beaker containing a mixture of sulphuric acid and water. 'Stuff of *that* kind is a mixture of H_2O and H_2SO_4', said pointing to the contents of the beaker, is, it seems to us, necessary if 'stuff of that kind is H_2O' is. Similarly, pointing at a liger, 'things of that kind are the offspring of a male lion and a female tiger' would seem to be necessary (again, if 'stuff of that kind is H_2O' is). But mixtures and biological categories are not, on Ellis's view, natural kinds: they do not have essences, and they do not carve nature at its natural joints.

The upshot is that Ellis does not have a story about what distinguishes *de re* necessity from other kinds of necessity—and in particular analyticity—that is consistent with his own view about what kinds of thing count as natural kinds. He cannot consistently say that no analytic truth is a truth about *de re* necessity—since that effectively rules out any kind for which there is no available term that meets Kripke's conditions on baptism from counting as a natural kind (as in 'ununbium is the element with atomic number 112'). And he cannot consistently say that the relevant distinction can be drawn by appealing to those kinds that can be referred to by ostensive definition and those that cannot, since many non-natural (in Ellis's sense of 'natural') kinds—such as mixtures of H_2O and H_2SO_4, and ligers—fulfil that criterion, if natural kinds do.

5. DIAGNOSIS

Our diagnosis of Ellis's predicament is that he is operating with two different conceptions of 'natural kind'. Early on in his book, Ellis appears to distance himself from Kripke: 'my concerns', he says, 'are different from Kripke's. Kripke's essentialism was developed in relation to theories of reference and identity. Scientific essentialism derives from an examination of the scientific practice of theoretical identification . . . Scientific essentialism is primarily concerned with the question of what makes a thing the kind of thing that it is, and so display the manifest properties and behavior it does. That is, real essences for us must not only be identifying, they must be explanatory' (2001: 54).

Ellis thus aligns himself with a more Lockean conception of natural kinds:

> The scientific task is to discover what makes a thing the kind of thing that it is and hence to explain why it behaves or has the properties it has. The scientific version of essentialism is therefore less concerned with questions of identity, and more with questions of explanation, than is the classical essentialism of Aristotle or the new essentialism of Kripke. Its closest historical predecessor is the kind of essentialism described by Locke. For Locke too was concerned with the question of what makes a thing the kind of thing that it is. He thought that if only we knew this, we should be able to explain why it has the manifest properties it has and behaves as it does. (2001: 55)

On the one hand, then, Ellis clearly wants to say that the natural kinds are those that carve nature at its joints, and that an investigation into the 'essences' of natural kinds is an investigation into the underlying natures of things that explain why they behave in the way that they do. On this conception of natural kinds, one is likely to find the natural kinds in the fundamental particles, the periodic table, and so on—that is, to the scientific practice of classifying on the basis of underlying nature, and explaining overt behaviour in terms of that underlying nature. But—crucially—such a conception of natural kinds requires no commitment to essentialism, or at least no commitment to Ellis's brand of essentialism.[8] One can quite happily hold that scientific classification aims to uncover the joints in nature, and that doing so is apt to lead to fruitful explanations and predictions of manifest behaviour, without being committed to any kind of 'real' necessity. And one can thus hold that 'phosphorus trichloride' picks out a natural kind, while 'raven-or-writing-desk' does not, without being an essentialist.[9]

On the other hand, Ellis is up to his neck in Kripkean commitments: when he talks about the 'scientific practice of theoretical identification', for example, it is Kripkean theoretical identification—the kind that generates *a posteriori* necessity—that he has in mind. 'Ours is not an *a priori* essentialism', he says on the previous page. 'We think that the laws concerning the behavior of the most fundamental kinds of things in nature are *a posteriori* necessary' (2001: 53). And, as we have seen, his account of the distinction between *de re* necessity—the defining feature of essentialism—and mere analytic necessity is driven by Kripkean intuitions concerning our ability to ostend a kind of stuff without knowing its underlying nature. And Ellis's Kripkean commitments are absolutely vital to the essentialist enterprise, since, as we said earlier, without them we can happily endorse the view that there are natural kinds while eschewing essences and denying that the laws of nature are metaphysically necessary.

However—to return to our basic point—there is simply no *a priori* reason to think that the notion of 'natural kind' deployed by Kripke is coextensive with the notion of 'natural kind' that focuses on carving nature at its joints and the explanation of behaviour in terms of underlying nature. Such a radical extension of the Kripkean conception of natural kinds needs arguing for; Ellis, it seems to us, simply takes the extension for granted, because he fails to disambiguate between two quite different conceptions of 'natural kind'.

The aim of this chapter has not been to show that there is anything wrong with essentialism as a metaphysical doctrine. Our aim has rather been to show that Ellis (and by extension other philosophers who extend the remit of the necessary *a posteriori* without argument) is not entitled to assume that truths about essences (if there are such things) are knowable only *a posteriori*. Moreover, in Ellis's case at least (although we believe the case can be made for other philosophers too), we have argued that that assumption is simply false. The temptation to extend Kripke's category of the necessary *a posteriori* is understandable, for it is precisely this extension that allows scientific essentialists to uphold the two fundamental tenets of the view: first, that the world is not, at bottom, merely a Humean mosaic of matters of particular fact over which the laws of nature generalize, but rather a world in which the laws of nature are 'immanent'—part of the fundamental fabric of reality; and, second, that those laws are genuinely *discovered* to be true (and hence, given the nature of laws, metaphysically necessary). Our remit here has not been to show that these two theses are incompatible. It has merely been to show that appeal to the Kripkean story about natural kinds will not do the job.[10]

NOTES

1. Other philosophers who have similarly failed to provide convincing (or indeed any) arguments for their claims about the necessary *a posteriori* status of truths that fall outside the narrow scope of Kripke's own argument include Alexander Bird (2007) and Adrian Heathcote and David Armstrong (1991). Bird, like Ellis, holds that the laws of nature (conceived by Bird as propositions that lay bare the underlying essences of fundamental dispositions) are necessary *a posteriori*; Heathcote and Armstrong hold that the relation of nomic necessitation is, as a matter of metaphysical necessity, identical with the causal relation, but that this identity can only be known *a posteriori*. We cannot substantiate the claim that the relevant arguments are lacking here, but we invite sceptical readers to attempt to locate them for themselves.
2. In the case of Heathcote and Armstrong's claim that 'causation is nomic necessitation' is necessary *a posteriori*, it is less clear whether they intend nomic necessitation to be the 'essence' of causation (this is suggested by their claim that 'investigation shows that causal sequences are essentially nomic' [1991: 67]) or whether they are thinking of 'causation is nomic necessitation' to be a genuine identity statement, analogous to 'Hesperus is Phosphorus'.
3. We, like Mellor (1977), Zemach (1976), and, more recently, Needham (2000), object to the drastic oversimplification of the natural kind essence claims found in Kripke, Putnam, and much of the subsequent literature. Our concern here, however, is not with essentialism itself, but rather that any potential essence claim, expressed on the right-hand side of a theoretical identity, must include so must empirical information that it is *obviously* an attributive definite description. Consider H_2O molecules: since H_2O molecules are

made of hydrogen and oxygen atoms, and each type of atom has three stable isotopes (atoms with the same atomic number but a different atomic mass), there are 18 constitutional variations of an H_2O molecule. Furthermore, in order for collections to actually be molecules, the constituent atoms must bond together. Molecules of H_2O have a polar covalent bond: polar because of the way that the charge is distributed within the molecule, and covalent because each bond (of which there are two) is between two atoms maintained by a shared pair of electrons. The molecule also has a particular geometry: as the oxygen atom has six electrons, and only two have those are being shared with the two hydrogen atoms (which also contribute a single electron each to the two covalent bonds), the molecular geometry of an H_2O molecule is tetrahedral. However, although the standard bond angle of a tetrahedral molecule is 109°, and the bond angle in an H_2O molecule is 104.5°, owing to the mutual repulsion from the lone pairs of electrons.

4. We have borrowed the term ‘descriptor’ from Connelly et al. (about which more below): ‘The primary aim of chemical nomenclature’, it says, ‘is to provide methodology for assigning descriptors (names and formulae) to chemical species so that they can be identified without ambiguity, thereby facilitating communication’ (Connelly et al. 2005: 3). We shall have more to say about chemical nomenclature in the next section.
5. A temporary designator is a type of descriptor. Broadly speaking, a descriptor is a name or a chemical formula that unambiguously refers to a chemical kind in virtue of encoding decisive descriptive information. A temporary designator is a descriptor that has been introduced, as the name suggests, temporarily, since the element in question has yet to receive a permanent name and/or symbol.
6. Since element 112 has now been confirmed, the temporary designator will be replaced by a name suggested by Hoffman and approved by IUPAC. At that point—we would argue—the name for the element with 112 protons (‘hoffmanium’, perhaps) will stipulatively define the kind, rather than being a descriptor (‘ununbium’).
7. Thanks to Jessica Pfeifer for pressing this objection, and to Josh Parsons for suggesting the response.
8. See Leary 2009: 273–92.
9. Of course, one needs to give an account of the difference between the natural and the non-natural kinds, but there are available theories to choose from which do not require a commitment to essentialism. See, for example, Dupré (1986: 441–7, and 1995), Mellor (1977: 299–312) and Mumford (2005: 420–36).
10. Thanks to audiences at the metaphysics of science conferences at Grenoble in December 2008 and Melbourne in July 2009, in particular Alexander Bird and Brian Ellis, for helpful comments and criticism; also to Robin Hendry.

REFERENCES

Armstrong, D. M. (1983) *What Is a Law of Nature?* Cambridge: Cambridge University Press.

Bird, A. (2007) *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.

Connelly, N. G., Damhus, T., Hartshorn, R. M., and Hutton, A. T. (2005) *Nomenclature of Inorganic Chemistry: IUPAC Recommendations 2005*. Cambridge: Royal Society of Chemistry.

Donnellan, K. (2008) 'Reference and definite description', in A. P. Martinich (ed.), *The Philosophy of Language*. Oxford: Oxford University Press.

Dupré, J. (1986) 'Sex, gender and essence', in P. A. French, T. E. Uehling, and H. K. Wettstein (eds), *Midwest Studies in Philosophy XI: Studies in Essentialism*. Minneapolis: University of Minnesota Press.

———. (1995) *The Disorder of Things*. Cambridge, MA: Harvard University Press.

Ellis, B. (2001) *Scientific Essentialism*. Cambridge: Cambridge University Press.

Heathcote, A. and Armstrong, D. M. (1991) 'Causes and laws', *Noûs*, 25: 63–73.

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

LaPorte, J. (2004) *Natural Kinds and Conceptual Change*. New York: Cambridge University Press.

Leary, N. (2009) 'How essentialists misunderstand Locke', *History of Philosophy Quarterly*, 26: 273–92.

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

Mumford (2005) 'Kinds, essences, powers', *Ratio*, 18: 420–36.

Needham, P. (2000) 'What is water?', *Analysis*, 60: 13–21.

Putnam, H. (1973) 'Meaning and reference', *The Journal of Philosophy*, 70: 699–711.

Salmon, N. (2005) *Reference and Essence*. New York: Prometheus Books.

Soames, S. (2002) *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity*. New York: Oxford University Press.

Zemach, E. (1976) 'Putnam's theory on the reference of substance terms', *The Journal of Philosophy*, 73: 116–27.

10 Crosscutting Natural Kinds and the Hierarchy Thesis

Emma Tobin

1. INTRODUCTION

It is often argued that natural kinds form a hierarchy: if any two kinds overlap, then one must be subsumed under the other as a subkind (Kuhn 2000b: 228–52; Ellis 2001: 67–76, 97–100, 161–70). For example, if crocodiles and humans are classified as *vertebrates*, and humans are classified together with gorillas as *mammals*, then gorillas and crocodiles should also be classified together under one of the categories (in this case vertebrates). Thus, the kind mammal can be subsumed as a subkind of the kind vertebrate. There are, however, many examples in both biology and chemistry of crosscutting kinds that do not form such simplistic nested hierarchies. This chapter examines whether the existence of such crosscutting categories in scientific taxonomy can be reconciled with the hierarchy thesis (HT).

There are several cases of crosscutting categories in biological taxonomy. Humans and dogs are classified together as *mammals*, and dogs and crocodiles are classified together as *quadrupeds*. However, crocodiles and humans cannot be classified together as either mammals or quadrupeds. Given the hierarchy thesis, the quadrupeds would have to be rejected as a legitimate kind category (Khalidi 1998: 102). Such cases of crosscutting abound in biological taxonomy and are often taken as evidence that species should be construed as individuals rather than kinds (Kitcher 1984; Dupré 1993; Ereshefsky 1992).

In contrast, chemical kinds are taken to be paradigmatic examples of natural kinds. For example, the elements magnesium (Mg) and promethium (Pm) are classified together as *metals*[1]. The elements lanthanum (La) and promethium (Pm) are classified together as *lanthanides*. Therefore, in accordance with the hierarchy thesis, magnesium and lanthanum must also be classified together as members of one or other of these kinds. This is the case, since the kind lanthanides is a subkind of the kind *metals* (e.g. all lanthanides are *metals*).

In theory, higher-level chemical kinds (e.g. molecules and macromolecules) should also form simple hierarchies. Such kinds are mereologically related as wholes to their composing elements. Thus, classification of the composing elements should suffice for classification of the molecular or macromolecular whole. If this were the case, then chemical taxonomy would appear to support the hierarchy thesis at all levels of classification.

However, cases of crosscutting categories abound in chemical taxonomy too. For example, albumin and renin can be classified together as *proteins*. Renin and the hairpin ribozyme can be classified together as *enzymes*. However, the hairpin ribozyme and albumin cannot be classified together as either enzymes or proteins. Enzymes are not a subkind of the kind *proteins* and proteins are not a subkind of the kind *enzymes*[2]. Such cases of crosscutting make it impossible to provide a neat hierarchical account of these kinds. Nevertheless, the classifications they provide would appear to be more than merely conventional. Metaphysical accounts of natural kinds must take such cross-classifications into account.

Some philosophers argue that because equally legitimate categories crosscut each other, then the ideal of a taxonomic hierarchy of natural kinds ought to be rejected (Khalidi 1998). Hacking (2007) agrees that the hierarchy thesis should be rejected, but furthermore argues that the distinction between all kind categories is merely conventional. Others argue that crosscutting categories lend evidential support for pluralism (Kitcher 1984; Dupré 1993; Ereshefsky 1992).[3] On the other hand, natural kind realists agree that crosscutting categories cannot delineate real natural kinds, but argue that some categories are categorically distinct and thus delineate real natural kinds (Ellis 2001: 67–76).

> The consensus amongst all of these views is that cases of crosscutting seriously jeopardize the view that natural kinds can be construed realistically. This paper supports the claim made by Khalidi (1998) and Dupré (1993) that crosscutting categories entail the rejection of the hierarchy thesis. Nevertheless, metaphysical accounts of natural kinds must allow for crosscutting categories. Crosscutting categories in science entail the rejection of the hierarchy thesis. *Pace* Hacking (2007), the rejection of the hierarchy thesis does not entail conventionalism about natural kinds.

2. THE HIERARCHY THESES AND CROSSCUTTING CATEGORIES

According to the hierarchy thesis, natural kinds form a nested hierarchy. If crocodiles and humans are classified as *vertebrates*, and humans are classified together with gorillas as *mammals*, then gorillas and crocodiles should also be classified together under one of the categories (in this case vertebrates). Thus, the kind mammal can be subsumed as a subkind of the kind vertebrate.

Historically, the Linnean system of classification in biology, which was originally formulated by Linneaus in his *Systema Naturae* (1735), provides a clear example of the ideal taxonomy envisaged by the hierarchy thesis.[4] In the Linnean system of classification, organisms are grouped into species, species into higher-level genera, genera into families, families into orders, orders into classes, classes into phyla, and phyla into kingdoms. For example, the tiger (*Panthera tigris*) can be subsumed under the genus *Panthera* (the panthers). The genus *Panthera* can be subsumed under the family *Felidae* (the cats). The family *Felidae* can be subsumed under the order *Carnivora* (the carnivores), and so on upwards until we reach the Kingdom *Animalia* (the animals).

Similarly, in chemistry, the periodic table of the elements might be viewed as an ideal hierarchical system of classification. The chemical elements as divided in the periodic table are supposed to reflect natural divisions between the elements in nature. In this respect, chemical kinds are considered to be categorically distinct, insofar as each element is individuated in terms of its atomic number (the number of protons in its nucleus). Moreover, grouping such elements together (e.g. the alkali metals in group 1 of the periodic table) involves classifying them according to the homologous behaviour of all the elements in that group (e.g. high reactivity, forming soluble oxides of the form X_2O, reactivity with water to form alkali solutions). This behaviour is a direct result of patterns in the electron configuration of each element in the group. For example, the element sodium can be subsumed under the kind *alkali metals*, which can be subsumed under the more general kind *metals*.

There are several motivations for the hierarchy thesis. The first motivation to consider is a semantic one about the role of natural kinds in language. Kuhn's chief motivation for endorsing the hierarchical structure of natural kinds was to show how the problem of incommensurability arises (Kuhn 2000a, b). The relationship between kinds in the hierarchy would provide a sufficient restriction on kind terms thereby avoiding incommensurability. Kuhn developed the no-overlap principle: real natural kinds would not overlap unless they were related one to the other as genus to species. The no-overlap principle precludes cross-classification of objects into different kinds within a theory's taxonomy. For example, there is no gold that is also silver, and that is what makes the terms 'silver' and 'gold' kind terms.

Scientific revolutions break the no-overlap principle (Kuhn 2000a: 92–6) and theories separated by a revolution cross-classify the same things into mutually exclusive sets of kinds. For example, according to Ptolemy's geocentric theory the sun is classified as a planet, which orbits the earth, while according to the Copernican heliocentric theory the sun is classified as a star that is orbited by the earth. These conflicting classifications are mutually exclusive with another. This would result in conflicting expectations about the sun depending on which taxonomy is used. The two taxonomies are thus incommensurable.

The second motivation is a naturalistic one, namely, the desire for a single unambiguous system of classification for picking out the real divisions in the natural world. Ellis (2001: 67–72) defends the hierarchical view of natural kinds on the basis that the natural sciences discover natural kinds as the primary objects of their investigation. Since science seems to deliver a hierarchy, our best ontology should also be hierarchically structured. He claims that an account of natural kinds needs to be given if we are to construct an ontology that is adequate for the natural world.[5] The ontology he puts forward is one that involves hierarchies of objects of increasingly complex kinds.[6]

Distinct hierarchy theses need to be carefully distinguished. The first hierarchy thesis that can be distinguished involves the overlapping of taxa. It can be delineated in the following way:

> **H1:** Natural kinds form a hierarchy: if any two kinds overlap then one must be subsumed under the other as a subkind (Thomason 1969, Kuhn 2000b). Let $<$ be the relation of species to genus or genus to higher order taxa. No natural kinds a and b of a taxonomic system overlap unless $a < b$ or $b < a$ or $a = b$. Thomason (1969: 98)

Examples of this kind of subsumption abound in biological classification. An example is provided by the gorilla/human/crocodile case described earlier. Similarly, there are examples of this kind of subsumption in chemical classification. For instance, in accordance with the hierarchy thesis, magnesium and lanthanum must also be classified together as members of a kind. This is the case, since the kind lanthanides is a subkind of the kind metals (e.g. all lanthanides are *metals*).

However, as we have already seen there are numerous examples of categories, in both biological and chemical classification, which do not form the kind of simplistic nested hierarchies that we have seen earlier. For example, humans and dogs are classified together as *mammals*. Dogs and crocodiles are classified together as *quadrupeds*. Nevertheless, humans and crocodiles cannot be classified together as either *mammals* or *quadrupeds*. Mammals are not a subkind of quadrupeds and quadrupeds are not a subkind of mammals (Khalidi 1998: 102). There are several responses available to the advocates of the hierarchy thesis. Firstly, one might argue that the category *quadruped*, based on means of locomotion, is not a natural kind category anyway. Instead, one could argue that the real kind category to be considered is *tetrapod*. But, since humans, although not quadrupeds are *tetrapods*, the hierarchy thesis can be maintained. Since all mammals are tetrapods and all quadrupeds are tetrapods, then the categories mammals and quadrupeds can be subsumed under the category tetrapod.

Thus, defenders of the view that natural kinds have a hierarchical structure might respond by modifying the hierarchy thesis accordingly. Overlapping occurs, but one is not subsumable under the other; rather, there is some common genus under which the two overlapping kinds can be subsumed.

In other words, the kinds *mammals* and *quadrupeds* are subkinds of the kind *tetrapod*.

> **H2:** When two kinds overlap and one is not a subkind of the other, then both kinds have a common genus. If two species *a* and *b* of a taxonomic system overlap then there is some kind *c*, where $a < c$ & $b < c$.[7]

Ellis advocates the latter thesis:

> the membership of two distinct natural kinds cannot overlap, so that each includes some, but not all, of the other, unless there is some broader genus that includes both kinds as species. (Ellis 2001: 20)

Consider, however, the following example of crosscutting between the chemical kinds *proteins* and *enzymes*. This example cannot be accommodated by either of the two hierarchy theses. Albumin and renin are classified together as *proteins*. Albumin is an umbrella term for any kind of water-soluble protein. An example is serum albumin, which is the largest plasma protein in humans and is composed of 584 amino acids.[8] Renin, secreted in the kidneys, is a protein composed of a sequence of 406 amino acids. Renin is also classified as an *enzyme*[9]. Likewise, a ribozyme, such as the hairpin ribozyme, is an autocatalytic RNA molecule that can also be classed as an *enzyme*. Thus, renin and the hairpin ribozyme can be grouped together as *enzymes*.

Until the 1980s, all known enzymes were thought to be proteins, until Thomas C. Cech discovered that RNA molecules could themselves catalyze chemical reactions.[10] These catalytic RNA molecules are called Ribozymes. Until Cech's discovery, RNA (ribonucleic acid) was considered to be merely a copy of the instructions given in DNA. RNA was only a messenger that could direct protein synthesis. Cech discovered that RNA could itself fold into different shapes and in so doing could catalyze its own biochemical reactions. This functional role was previously thought to be restricted to protein enzymes. Therefore, even though renin and the hairpin ribozyme can be classified together as *enzymes* and renin and albumin can be classified together as *proteins*, albumin and the hairpin ribozyme are not classified together as either *enzymes* or *proteins*, since not all enzymes are proteins and not all proteins are enzymes.

The first hierarchy thesis (H1) cannot accommodate this example, because proteins and enzymes overlap, but proteins are not a subkind of enzymes and enzymes are not a subkind of proteins. Equally, the second hierarchy thesis (H2) cannot accommodate this example because there is no common higher order genus, under which proteins and enzymes can be subsumed.

One possible response is to claim that proteins and enzymes belong to the higher genus *biomolecule*. A biomolecule is any organic molecule (e.g. metabolites) or macromolecule (e.g. proteins and enzymes) produced by a living organism. However, members of the kind biomolecule are being grouped together solely in virtue of being produced by a living organism. Grouping them together under this category masks important differences and similarities between the members, which is made clear by the fact that they crosscut each other.

Alternatively, hierarchy theorists may claim that subsumption is not as scientifically informative as looking for a common underlying microstructure in these cases. In other words, there is a common underlying kind of which these kinds are composed. It might be argued, for example, that proteins and enzymes have the same underlying structure, because they are essentially composed of DNA. This would lead to the following hierarchy thesis:

> H3: When two kinds overlap and one is not a subkind of the other, then both kinds have a common underlying structure. If two species *a* and *b* of a taxonomic system overlap then there is some underlying kind *c* of which *a* and *b* are composed.

H3 might seem like a special case of H2 where the common genus is some microstructural kind of which both overlapping kinds are composed. For example, we might expect that the microstructural kind DNA is the kind c, of which proteins and enzymes are composed. However, this is not straightforwardly the case. The underlying mechanisms involved are different to each other. Firstly, the hairpin ribozyme is a self-splicing RNA molecule, in other words RNA having a catalytic effect on RNA. Alternatively, in the case of renin and albumin the underlying mechanism involved is DNA directing its own replication.[11] Clearly, the underlying structure is not the same.

It might be argued that the kind *nucleic acid* is the underlying kind of which both enzymes and proteins are composed. Thus, there is a common underlying structure involved. We could certainly subsume the kinds enzyme and protein under the kind nucleic acid and thus claim that H3 can be supported. However, to do so would be misleading in that the kind nucleic acid masks distinct structural differences between RNA and DNA. These differences are indicated by the fact of higher level crosscutting kinds. Thus, to subsume them under a homogeneous grouping would be ontologically misleading.

It is also worth making the point that reduction is not straightforward in these cases. Reduction is precluded by the fact that categories such as albumin and ribozymes are determinable categories. There are several determinates of these determinables. For example, human blood serum and ovalbumin (egg white) are both determinates of the determinable albumin. Importantly, they have significantly different underlying structures. The ovalbumin protein is made up of 385 amino acids, while the serum albumin protein is made up of 584 amino acids.

Similarly, there are many different determinates of ribozymes, because there are different examples of catalytic RNA molecules. Some examples of ribozymes are viroids, which are RNA molecules that infect plant cells.[12] These can be classified together with Ribonuclease P, which cleaves tRNA. Ribonuclease P is a ubiquitous enzyme present in many different kinds of cells such as bacterial cells like *E. coli*, eukaryotic nuclei, mitochondria, and chloroplasts (Gopalan, Vioque, and Altman 2002). Classification simply in terms of underlying chemical structure would omit important scientific classifications, namely the fact that albumin refers to any water-soluble protein and that ribozymes refers to any catalytic RNA molecules.

Thus, we ought to reject H3 for two reasons. Firstly, as we have seen in some cases of crosscutting there is not one underlying kind of which the crosscutting categories are composed. Some cases of crosscutting are ontologically significant in that the underlying kinds are in fact distinct. Thus, crosscutting higher order categories indicate real differences in kind, in that the failure to provide a simple nested hierarchy indicates a real difference in underlying structure. Secondly, the prospects for reducing higher level kinds is unpromising, since higher level kinds are determinables whose determinates are microstructurally different (e.g. albumin and ribozymes).

In this section, three hierarchy theses were distinguished. However, a consideration of one example of crosscutting has revealed that none of the hierarchy theses is satisfactory. H1 claimed that if any two kinds overlap, then one must be subsumed under the other as a subkind. However, it is clear that proteins and enzymes overlap, but proteins are not a subkind of enzymes and enzymes are not a subkind of proteins. H2 claimed that when two kinds overlap and one is not a subkind of the other, then both kinds have a common genus. However, there is no common higher order genus, under which proteins and enzymes can be subsumed. H3 claimed that when two kinds overlap and one is not a subkind of the other, then both kinds have a common underlying structure. We have seen in the ribozymes/proteins case that there is not one underlying kind of which the crosscutting categories are composed. Equally, the reduction of higher level determinable kinds is unpromising. Thus, a closer examination of the hierarchy thesis has revealed that there is no satisfactory formulation that allows for cases of crosscutting categories.

Khalidi (1998: 50) has already pointed out that crosscutting categories make the ideal of a single hierarchical taxonomy of natural kinds look unpromising, though he admits that a different articulation of natural kinds could nevertheless be possible. I wish to argue, *pace* Hacking (2007), that the rejection of the hierarchy thesis does not entail conventionalism about natural kinds. Indeed, realism about natural kinds can be articulated without the hierarchy thesis, to allow for such cases of crosscutting categories.

3. CROSSCUTTING AND THE METAPHYSICS OF NATURAL KINDS

What can be concluded about the metaphysics of natural kinds from the rejection of the hierarchy thesis? One possibility is to claim that natural kinds should not be construed realistically at all; in other words, all natural kinds are conventional. Hacking (2007) motivates this kind of view when he argues that the desire to accommodate natural kind categories into a simplistic tree-like hierarchy, organized around subsumption principles, is futile.

> The periodic table is a permanent refutation of the idea that natural kinds have to be organized into a hierarchy. There are obvious genera and species within the table, for example the halogens form a genus of which chlorine and iodine are species. But the structure is not a simple hierarchic set of nested sets. (Hacking, 2007: 214)

Hacking argues that higher level crosscutting categories should not be construed realistically as natural kinds. In fact, he makes the somewhat stronger claim that 'the concept of a natural kind, which began in a promising way and has taught us many things, is now obsolete.' (Hacking, 2007: 205). For Hacking, all natural kind classifications are conventional and are not 'real Kinds' in Mill's sense.[13]

From considerations about cases of crosscutting, some realists have argued for a less radical conclusion. Ellis (2001) agrees with Hacking that crosscutting kinds are not real natural kinds, because they are not categorically distinct. Nevertheless, Ellis allows that some kinds are real natural kinds, namely those kinds that are categorically distinct. For example, the property kind *mass*, which is categorically distinct from charge and from every other quantitative property, is a real natural kind for Ellis (2001: 71). In contrast, biological kinds (e.g. species) are not.[14] He states:

> Because of the messiness of biological kinds, and in order to develop a theory of natural kinds adequate for the purposes of ontology, I have broken with the tradition of using biological examples, and taken the various kinds of fundamental particles, fields, atoms and molecules as paradigms. (2001: 170)

Ellis argues that chemical kinds are categorically distinct natural kind categories.[15] In fact, he takes chemical kinds to provide a paradigmatic example of categorical distinctness. This is also the case for higher level chemical kinds since they are hierarchically related to their composing parts. So, higher level chemical kinds (molecules and macromolecules) are simply composed of lower level chemical kinds (e.g. elements) in a particular kind of chemical reaction.[16]

Though Ellis and Hacking have opposing views about natural kinds, they both agree that if there are any real natural kinds, then they must be categorically distinct. Thus, any two categories that crosscut each other are not natural kind categories. In the previous section, a case of crosscutting in biochemistry has been outlined. Should we conclude from this case of crosscutting that the lines between kinds of macromolecules are just as arbitrary as those that are drawn between species? If so, then we must conclude that macromolecules are not natural kinds.

This conclusion would be premature.[17] One might argue that there are genuine systems of classification (i.e. ones that reflect natural similarities and differences) that are not classifications by natural kinds. Thus, we could argue that Ellis is correct about which natural kinds there are, and that they obey either H1 or H2. But in addition to the natural kind classifications there are other 'natural' classifications (e.g. macromolecular classification), which don't conform to any of the hierarchy theses outlined in the previous section. In fact, Ellis (2001: 170) concedes that although biological kinds are not natural kinds, they are nevertheless real classifications. Thus, he agrees that conventionalism certainly does not follow directly from the availability of crosscutting categories.

However, I think there is evidence for an even stronger claim, namely that macromolecular classification is a system of natural kind classification, but which nevertheless does not form simple nested hierarchies. There are two important distinctions between crosscutting cases in biology and biochemistry. Firstly, we need to distinguish between cases of crosscutting in biology (e.g. species) and cases of crosscutting in biochemistry (e.g. macromolecules). Crosscutting is more vicious in the case of species. The examples that were considered in §1 were examples of *intrataxonomic* crosscutting: in other words, crosscutting categories within a single taxonomic system. So, taking as an example the Linnean system of biological classification or the periodic table of elements in chemistry, cases of crosscutting within these systems were considered. However, what makes the species problem more vicious is that there are also cases of *intertaxonomic* crosscutting.[18] Intertaxonomic crosscutting occurs when there are different species concepts, across different taxonomic systems, which delineate species in radically different ways.[19]

An example of intertaxonomic crosscutting can be seen from the fact that different species concepts draw distinct lines between species. According to morphological species concepts, species ought to be individuated by means of similarities such as colour, shape, pattern, and bone structure. Alternatively, according to interbreeding species concepts, to be members of the same species organisms must be capable of interbreeding. Depending on which species concept is used, the lines between species is drawn differently. For example, western meadowlarks (*Sturnella neglecta*) and eastern meadowlarks (*Sturnella magna*) are almost identical to one another; thus according to the morphological species concept they are members of the same species. However, they do not interbreed, so according to the interbreeding species concept, they are not members of the same species.

There is no equivalent to intertaxonomic crosscutting in chemistry: the periodic table of elements is considered to be a stable taxonomic system. Even though there are certain disagreements about how it is presented, the core elements arranged according to their atomic number remain the same (see Mazurs 1974). Therefore, species of chemical element are not delineated differently according to different chemical models.

There is a second distinction between cases of crosscutting in biology and chemistry. In the case of species, crosscutting can occur because species are always evolving and thus can change into different kinds over time. For example, in the event of reproductive isolation a subpopulation can result in a new species. Reproductive isolation will split existing populations and the newly isolated subpopulations will belong to a new species. For example, according to the interbreeding species concept, species are individuated according to their ability to interbreed. In the event of reproductive isolation the boundaries between the species would at least for a time be vague. Geographical isolation would prevents interbreeding, but if these species were not geographically isolated, then they would still be capable of interbreeding. Over time, geographical isolation would result in the two species evolving differently and thus no interbreeding would be possible. Thus, there will a time period after reproductive isolation, where the lines between the two species will be difficult to draw, at least according to the interbreeding species concept.[20]

However, there are no analogous cases of crosscutting through time in chemistry. Consider cases of chemical transmutation. In cases of beta decay, an individual element changes its atomic number. The nucleus will have either one more proton or one less proton after decay. One and the same nucleus persists through this transformation. An individual nucleus can change from being identifiable *qua* carbon to being identified *qua* nitrogen at a later stage. An individual can retain its identity while undergoing a change of kind. Chemical elements do not evolve in the same way as biological species do. Thus, atomic number is accepted as the one and only chemical species concept for the elements.

The macromolecular cases of crosscutting that were considered in §1 are all cases of intrataxonomic crosscutting. Equally, these are not cases of crosscutting through time, where an individual or the members of a kind can change. Rather, they are cases where the kinds themselves do the crosscutting. The higher level kinds crosscut the underlying microstructural kinds, which compose them. We have seen earlier that the chemical kinds *proteins* and *enzymes* crosscut each other. Some enzymes are not proteins (e.g. ribozymes). But the underlying mechanisms involved are different to each other. Ribozymes, such as the hairpin ribozyme, are composed of self-splicing RNA molecules. Alternatively, in the case of proteins like renin and albumin the underlying mechanism involves DNA directing its own replication into messenger RNA.

Thus, in this case, the crosscutting higher order categories indicate real ontological differences in kind, in that the failure to provide a simple nested hierarchy indicates a real difference in underlying structure. To insist on the simplistic nested hierarchies involved in the hierarchy thesis would be to ignore the ontological significance of such crosscutting categories. Therefore, crosscutting categories in biochemistry are less vicious than crosscutting biological species. An acceptance of such crosscutting categories need not entail that the boundaries between them are arbitrary.

In conclusion, §1 of this paper argued for the rejection of the hierarchy thesis concerning natural kinds. Section 2 illustrated that theorists with radically different accounts of natural kinds agree that the acceptance of crosscutting categories should entail conventionalism about those categories. However, I argue that it would be premature to conclude that crosscutting categories are not natural kinds. A closer examination of some crosscutting kinds at the macromolecular level reveals that crosscutting can be ontologically significant; namely, where chemical kinds crosscut there is a real difference in the chemical microstructure. *Pace* Hacking (2007), the rejection of the hierarchy thesis need not entail conventionalism about natural kinds. Metaphysical accounts of natural kinds need to accommodate such crosscutting categories.[21]

NOTES

1. It should be noted that some other examples of the kind metals might not be so easily accommodated. For example, tin (Sn) has two forms (allotropes): 'white tin', which is a metal, and 'grey tin', which is more covalent in character and is a non-metal.
2. A detailed discussion of this example follows in §1.
3. Pluralism is the view that there are several different equally legitimate ways of carving nature into natural kinds. For example, in biology pluralists do not believe that there is a single correct species concept. They argue that there are a number of legitimate species concepts, which divide species according to different interests.
4. See Ereshefsky 2000 for a discussion.
5. Hacking (1991: 111) refers to this as the uniqueness principle: the claim that 'there is a unique best taxonomy in terms of natural kinds, that represents nature as it is and reflects the network of causal laws'. Hacking argues against the ideal of a complete exhaustive taxonomic framework.
6. See Ellis 2001: Ch. 2 for a discussion of his six-category ontology. Ellis distinguishes between substantive kinds (natural kinds of objects), dynamic kinds (natural kinds of processes), and property kinds (natural properties). For Ellis, the existence of a hierarchy of natural kinds is a corollary of the existence of a natural hierarchy of causal powers and other dispositional properties.
7. < is to be understood as the relation of subsumption.
8. Serum albumin transports essential fatty acids from fat tissue to muscle tissue. It also has a role in the regulation of osmosis, and helps to transport substances (e.g. hormones) through the blood.

9. Renin regulates mean arterial blood pressure.

10. See Cech 1989 for Cech's Nobel Prize lecture, which discusses autocatalytic RNA molecules, reprinted in Cech (1990).

11. Each gene is independently copied into a nucleic acid called messenger RNA (mRNA).

12. See Collins et al. 1998 for a discussion.

13. For Mill (1843: 170–8), real Kinds are distinguished by the multitudinous properties that actually characterise them, rather than the few choice properties that essentialists might deem constitutive of the kind. He states: 'every class which is a real Kind, that is, which is distinguishable from all other classes by an indeterminate multitude of properties not derivable from another, is either a genus or a species' (1843: 171).

14. The view that species are natural kinds has come under much scrutiny and in recent times has become difficult to sustain. Biologists have put forward a variety of species concepts which disagree on how species are individuated, and philosophers have also questioned whether the lack of consensus amongst biologists reflects ontological divisions between species themselves; namely species pluralism.

15. See Ellis 2001: 161 for some examples.

16. Ellis (2001) does not explicitly state that periodic groups in the periodic table represent natural classifications. We saw earlier that Hacking (2007) argues that groups like the halogens are not real natural kinds. Presumably, Ellis would state that periodic groups are natural classifications that rely upon essential properties of the composing members. For example, all halogens have the same patterns in their electron configuration, and are highly reactive because of their atoms being highly electronegative.

17. There are several possibilities for realism about natural kinds, which do not require the full-blown variety of metaphysical realism that we find in theorists like Ellis (2001) and Lowe (2006). Dupre's (1993) promiscuous realism provides an example.

18. Of course, it could turn out that there is one correct species concept, according to which we ought to delineate species. Thus, whether intertaxanomic crosscutting is really the case is an open question.

19. There are many competing species concepts. Mayr's (1970) biological species concept, Darwin's (1859) morphological concept, and Hennig's (1968) cladistic concept provide just a few examples.

20. See Bird and Tobin 2008 for a discussion of how kind essentialism does not entail essentialism of kind membership. The point is made originally in LaPorte 1997, and also in his 2004.

21. I am grateful to Helen Beebee, Alexander Bird, Nigel Sabbarton-Leary, and Samir Okasha for comments on earlier drafts of this paper. I also wish to acknowledge the AHRC for financially supporting a period of postdoctoral research, during which this paper was written.

REFERENCES

Bird, A. and Tobin, E (2008) 'Natural Kinds', in E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* (winter 2008 edition). Online. Available HTTP: .

Cech, T. (1990) 'Self-splicing and enzymatic activity of an intervening sequence RNA from tetrahymena'. *Bioscience Reports*, 10: 239–61.

Collins, R. F., Gellatly, D. L., Sehgal, O. P., and Abouhaidar, M. G. (1998) 'Self-cleaving circular RNA associated with rice yellow mottle virus is the smallest viroid-like RNA', Virology, 241: 269–75.

Cracraft, J. (1983) 'Species concepts and speciation analysis', in R. Johnston (ed.), Current Ornithology, New York: Plenum Press, 159–87.
Darwin, C. (1859) The Origin of Species, G. Stade (ed.). New York: Barnes & Noble (2003).
Dupré, J. (1993) The Disorder of Things: Metaphysical Foundations of the Disunity of Science. Harvard: Harvard University Press.
Ellis, B. (2001) *Scientific Essentialism*. Cambridge: Cambridge University Press.
———. (2005) 'Physical realism', *Ratio* 18: 371–84.
Ereshefsky, M. (1992) 'Eliminative pluralism', Philosophy of Science 59: 671–90.
———. (2000) *The Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy*. Cambridge: Cambridge University Press.
Gopalan, V., Vioque, A., and Altman, S. (2002) 'RNase P: variations and uses', *Journal of Biochemistry*, 277: 6759–62.
Hacking, I. (1991a) 'A tradition of natural kinds', Philosophical Studies, 61: 109–26.
———. (2007) 'Natural kinds, rosy dawn, scholastic twilight', *Royal Institute of Philosophy Supplement*, 82: 203–39.
Hennig, W. (1966) *Phylogenetic Systematics*. Urbana: University of Illinois Press.
Khalidi, M. A. (1998) 'Natural kinds and crosscutting categories', *Journal of Philosophy, 95*: 33–50.
Kitcher, P. (1984) 'Species', Philosophy of Science, 51: 308–33.
Kuhn (2000a) 'The road since *Structure*', in T. S. Kuhn, eds J. Conant and J. Haugeland, *The Road Since Structure*, Chicago: Chicago University Press, 90–104.
———. (2000b) 'Afterwords', in T. S. Kuhn, eds J. Conant and J. Haugeland, *The Road Since Structure*, Chicago: Chicago University Press, 224–52.
LaPorte, J. (1996), 'Chemical kind term reference and the discovery of essence', Noûs, 30: 112–32.
———. (2004) Natural Kinds and Conceptual Change. Cambridge: Cambridge University Press.
Linnaeus (1735) *Systema Naturae*. Leiden: Theodorum Haak.
Lowe, E. J. (2006) *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford: Oxford University Press.
Mayr, E. (1969) Principles of Systematic Biology. New York: McGraw-Hill.
———. (1970), *Populations, Species and Evolution*. Cambridge: Harvard University Press.
Mazurs, E. G. (1974) *Graphical Representations of the Periodic System During One Hundred Years*. Tuscaloosa: University of Alabama Press.
Mill, J. S. (1843) *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence, and Methods of Scientific Investigation*. Honolulu: University Press of the Pacific, 2002.
Stenwall, R. (2007) 'Aspect kinds', *Boston Studies in the Philosophy of Science,* 252: 193–203.
Thomason, R. (1969) 'Species, determinates and natural kinds', *Noûs*, 3: 95–101.

11 From Constitutional Necessities to Causal Necessities

Jessica Wilson

Humeans and non-Humeans commonly and reasonably agree that there may be necessary connections ('necessities', for short) between entities that are identical—e.g. Hesperus and Phosphorus, water and H_2O—or merely partly distinct—e.g. sets and their individual members, fusions and their individual parts, instances of determinates and determinables, members of certain natural kinds and certain of their intrinsic properties, and (especially among physicalists) certain physical and mental states. Humeans maintain, however, as per 'Hume's Dictum', that there are no necessary connections between entities that are wholly distinct;[1] and in particular, no necessary *causal* connections between such entities (even when the background conditions requisite for causation are in place). The Humean's differential treatment appears principled, in reflecting the fact that commonly accepted necessary connections involve constitutional relations (involving, roughly, existential ontological dependence between certain entities), whereas wholly distinct entities (notably, causes and effects) do not constitute each other in this sense. I'll argue, however, that the appearance of principle is not genuine, as per the following conditional:

> *Constitutional → Causal*: If one accepts certain constitutional necessities, one should accept certain causal necessities.

This result provides needed leverage in assessing the two main frameworks in the metaphysics of science, treating natural kinds, causes, laws of nature, and the like. These frameworks differ primarily on whether Hume's Dictum is taken as a working constraint on theorizing; and it has proved difficult for either side to criticize the other without presupposing their preferred stance on the dictum, hence talking past one another. The arguments for *Constitutional → Causal* are based, however, on general and independent considerations about which facts in the world might plausibly warrant our beliefs in constitutional necessities involving (in particular) broadly scientific entities. The Humean can respond to these arguments, which reveal a deep tension in their view, only at attendant costs of implausibility and *ad hocery*. The non-Humean framework doesn't face any such tension between constitutional and causal necessities, however, and so in this respect comes out ahead.

1. CONSTITUTIONAL NECESSITY

1.1 Schaffer on Necessities of Identity vs. Causal Necessities

As motivation for the seemingly principled way in which Humeans may distinguish between constitutional and causal necessities, I want to first consider a proposal made by Schaffer (2004) that aims to provide a principled basis for accepting necessities of identity (a limiting case of constitutional necessity) while rejecting casual necessities.

Schaffer, like Hume, intuits that actual causes might have different effects; that, e.g., it is possible that like charges attract (again, even assuming that the background conditions requisite for actual cases of like charges' repelling are in place[2]). Hume assumed that conceivability was a sure guide to possibility; but in a post-Kripke climate, what force should intuitions of causal contingency have? After all, some (e.g. Wiggins 1965) intuited that identities were contingent—that Hesperus might not have been Phosphorus, or that water might not have been H_2O—but they were wrong;[3] hence such intuitions of contingency are either mistaken or misdescribed. Why shouldn't intuitions of causal contingency receive a similar treatment? Indeed, Shoemaker advocates a Kripkean redescription strategy, in defending his causal necessitarian view:

> Let the law be that strychnine in a certain dosage is fatal to human beings. We can grant it to be imaginable that ingesting vast amounts of what passes certain tests for being strychnine should fail to be fatal to what passes certain tests for being a human being, but deny that this amounts to imagining a human being surviving the ingestion of that much strychnine. (Shoemaker 1998: 62)

Schaffer rejects Shoemaker's suggestion on grounds that, unlike the case of identity, there is no compelling 'independent reason' for questioning intuitions of causal contingency: The Kripkean manoeuvre is compelling for water = H_2O because there is an identity, and identities are necessary (Kripke 1980: 97–105). Hence any conception of water being *XYZ* can only be an illusion. But the relation between (e.g.) charge and Coulomb's law is governance rather than identity, and hence no comparable compulsion to necessity exists. There is no independent reason for thinking that any misdescription is taking place. (Schaffer 2004: 218)

The sort of independent reason Schaffer has in mind (confirmed in p.c.) is one that (as per the talk of 'compulsion') is effectively a *proof* of the sort available for the case of identity.[4] In fact, Kripke does not rely on this proof, mentioned only in the second edition preface to *Naming and Necessity* (1980: 3), in arguing that intuitions of the contingency of identity are mistaken (he rather appeals to semantic intuitions exercised in consideration of hypothetical scenarios). But in any case it remains that there is a compelling, broadly logical proof of the necessity of identity, while there is not (we may assume) any such proof for the necessity of causal connection; so Schaffer's proposal does provide a principled basis for accepting identities of necessity while rejecting causal necessities.

On the other hand, Schaffer's proposal doesn't show how Humeans are generally within their rights to accept constitutional necessities, holding between entities that are not wholly distinct. As aforementioned, Humeans and non-Humeans reasonably agree that there may be necessary connections between entities that are not wholly distinct, including, e.g., sets and their individual members, fusions and their individual parts, instances of determinables and determinates, members of certain natural kinds and certain of their intrinsic properties, and (assuming non-reductive physicalism) certain physical and mental states. But the relations at issue in such necessities—set membership, parthood, the determinable/determinate relation, essential instantiation, realization—hold between non-identical entities, and so the proof of the necessity of identity does not apply to them.[5] Moreover, in none of these cases is there any compelling proof of their necessity. There is no proof from the axioms of set theory to the conclusion that sets are necessarily constituted by their members; set theory is silent on the modal individuation of sets. Nor is there any proof from the axioms of mereology to the conclusion that fusions are necessarily constituted by their parts. In other cases of commonly accepted constitutional necessity there is not even a set of candidate axioms characterizing the relation at issue from which we might try to derive the associated modal conclusion.

Rather, to the extent that anything resembling proof is involved in establishing such necessary connections, it is of one or other broadly non-logical variety, associated with metaphysical investigation into the entities or associated concepts at issue, as involving systematization of relevant beliefs, contemplation of hypothetical scenarios, transcendental arguments, or (perhaps concurrently with the previous approaches) inference to the best explanation of the various experiential, semantic, scientific, and philosophical facts. There are, of course, disputes both about how to implement these strategies and over the status of the results so gained (e.g. as *a priori* or not, as defeasible or not). These disputes aside, it remains that these sorts of 'proofs', which are the very lifeblood of philosophical theorizing, are, while often quite convincing, not the sort to *compel* assent, in anything like the way a broadly logical proof is able to do.[6]

To sum up: Schaffer's criterion of compelling proof warrants accepting necessities of identity while rejecting causal necessities, but fails to warrant acceptance of many constitutional necessities that Humeans and non-Humeans alike commonly and reasonably accept. In assessing this result, it is important to keep in mind that Humeans are not supposed to be in the business of denying constitutional necessities. Constitutional necessities involve necessary connections between entities that are not wholly distinct; hence—at least on the face of it—do not fall under the purview of Hume's Dictum, which dictum effectively characterizes the Humean's view. Relatedly, part of the indirect motivation for Hume's Dictum surely lies in its seeming that one can accept it without having to generally reject all or most commonly accepted constitutional necessities. As such, a better way for a Humean to resist taking intuitions of causal contingency to be mistaken would be to identify a principle that makes general room for constitutional necessities—including but not restricted to necessities of identity—while excluding causal and other necessities between wholly distinct entities, as Hume's Dictum requires.

1.2 Constitutional Necessity

A principle seemingly able to do this work isn't hard to find. As earlier, commonly accepted necessary connections between not-wholly-distinct, intrinsically characterized entities are cases of constitutional necessity (where identity is understood as a limiting case of constitution). The Humean may correspondingly suggest that constitutional necessity is the *only* sort of necessity between intrinsically characterized entities. In other words, they may endorse:

> *Constitutional necessity*: Intrinsically characterized entities are necessarily connected just in case (i) the entities are not wholly distinct, and (ii) at least one entity constitutes the other.

What is it for entities to be constitutionally connected? For present purposes it will suffice to work with a rough general account, according to which *a* constitutes *b* only if the existence of *a* is ontologically dependent on the existence of *b*, in that (at a minimum) *a*'s existence at a time *t* requires *b*'s existence at *t* (for short: *a* is existentially ontologically dependent on *b*). The rough account should be interpreted as allowing for variations on the theme, reflecting the ontological categories of the entities involved; so, for example, for properties the ontological dependence claims are best seen as involving (some or other) tokens of the types in question. Note that this understanding of constitution is neutral on the order of metaphysical priority, or relative fundamentality, of the entities involved: there is no presupposition that constituting entities are somehow more fundamental than (alternatively: serve as the truthmakers for claims about) constituted entities. For example, *being scarlet* is, on the present understanding, existentially ontologically dependent on *being red* (any instance of scarlet at a time *t* requires the existence of an instance of red at *t*); but on some accounts (instances of) determinates are more fundamental than (serve as the truthmakers for claims about instances of) determinables.[7]

Constitutional necessity appears well suited for the Humean's purposes. First, the principle makes room for Humean acceptance of some necessities besides necessities of identity. Second, the principle entails Hume's Dictum, since it rules out necessary connections between wholly distinct entities (as per condition (i)). Third, the principle does not countenance as necessary any and all connections between entities that are not wholly distinct: it is moreover required that the entities be constitutionally connected (as per condition (ii)).[8] As such, the principle doubly warrants the Humean's rejection of causal necessities, since, all parties agree, causes and effects are wholly distinct, and neither constitutes the other.[9]

Constitutional necessity, if true, would provide the Humean with a general, principled basis for accepting constitutional necessities while rejecting causal necessities; conversely, it is hard to see what other sort of principle would do the general trick. Endorsement of *constitutional necessity* (or some principle in the ballpark) thus appears to provide the best case for Humeans being able to have their constitutional necessities and Hume's Dictum, too. I will now argue, however, that a closer look at what facts might plausibly justify our beliefs in certain commonly accepted constitutional necessities indicates that *constitutional necessity* is false, for these facts presuppose that there are certain causal necessities.

Before continuing, I want to highlight a difference between my upcoming argument and Bird's (2001) argument aiming also to show that acceptance of certain constitutional necessities invokes commitment to certain causal necessities. Bird focuses on a necessity of identity, holding between salt and NaCl. Roughly, Bird argues that the truth of Coulomb's law is implicated both in the holding of the relevant ionic bonds between Na and Cl, and in the dissolution of these bonds in water; hence a world where salt exists is a world where Coulomb's law is true, hence a world where salt dissolves in water. Concerns may be raised about this argument (see, e.g., Beebee 2002) but for present purposes my concern is with the scope of its conclusion. Bird grants that fundamental causal laws are contingent, and aims only to show that if certain fundamental laws are in place, then certain non-fundamental laws are necessary. Consequently, Humeans can grant his conclusion, restricting their denial of necessary connections to those holding between wholly distinct entities that are sparsely, as well as intrinsically, characterized. This is, as it happens, how Lewis understands his thesis of Humean supervenience: only the fundamental constituents of the Humean mosaic are subject to Hume's Dictum; other entities receive different treatments but—as per the supervenience claim—both non-fundamental entities and laws necessarily follow once the contingent fundamental facts and laws are in place. Bird's conclusion thus doesn't provide leverage sufficient for a comparative assessment of Humean vs. non-Humean frameworks in the metaphysics of science. My argument aims to apply to constitutional necessities involving fundamental (or not obviously non-fundamental) as well as non-fundamental entities, with the associated consequent leverage for assessing Humeanism.

2. FROM CONSTITUTIONAL NECESSITIES TO CAUSAL NECESSITIES

My interest here is with competing Humean vs. non-Humean frameworks in the metaphysics of science, so let us consider three representative claims, expressing certain constitutional necessary connections between broadly scientific entities:

(N1) Necessarily, anything that is scarlet is red.
(N2) Necessarily, anything having a certain mean molecular kinetic energy (MMKE) has a certain temperature.
(N3) Necessarily, anything that is an electron is negatively charged.

N1–N3 express necessary connections between broadly scientific entities that are either identical or not wholly distinct, in being connected by way of the determinable-determinate relation (N1), identity or realization (N2), or natural kind constitution, where members of a natural kind essentially have an intrinsic property (N3).[10] In each of these cases it is plausible that the entities at issue stand in a constitutive relation: (instances of) scarlet existentially ontologically depend on (instances of) red; (instances of) temperature ontologically depend on (instances of) MMKE; (token) electrons existentially ontologically depend on (instances of) negative charge.

None of these claims is unassailable. Most saliently, one might reject N1 on Quinean grounds, N2 on anti-Kripkean grounds,[11] and N3 on anti-essentialist grounds. Such sceptical positions are rare, however, among those aiming to elucidate the nature of natural reality, and moreover are treated, by Humeans and non-Humeans alike, as orthogonal to acceptance or rejection of Hume's Dictum.[12] Hence it is that Humeans and non-Humeans alike are typically happy to accept analyticities such as N1 and realization or identity-based conditional claims such as N2. (Note also that N2 is the sort of claim that, when involving certain physical and mental states or properties, physicalists of both Humean and non-Humean persuasions accept.) One may wonder whether Humeans should balk at acceptance of N3, on grounds that negative charge is somehow fundamentally dispositional, in a way at odds with Hume's Dictum. Answer: no. Lowe (2006), for example, endorses Hume's Dictum as applied to causal necessities, and also endorses N3 ('all of a kind's nature may be essential to it in the simplest cases, such as that of the kind electron'); and Lewis's remark that 'there might have been altogether different laws of nature; and *instead of electrons and quarks, there might have been alien particles, without charge or mass or spin* but with alien physical properties that nothing in this world shares' (1986: 1, my emphasis) is clearly indicative of Lewis's acceptance of (a context in which) N3 is true. Ostensions aside, the deeper point is this: Humeans and non-Humeans agree that negative charge is a (good candidate for a) fundamental intrinsic property of fundamental entities; what they disagree about is whether or not such an intrinsic property is essentially dispositional. Each position is, on the face of it, compatible with maintaining that it is necessary that electrons have the intrinsic property in question. Relatedly, Humeans have good reason to insist (or at least hope) that they are within their rights to accept N3 while accepting Hume's Dictum, for as previously suggested, it is one thing to maintain that wholly distinct natural entities are entirely 'loose and separate', and quite another to maintain that fundamental natural entities have no essential intrinsic properties.[13]

What I now want to consider is: what facts about the entities at issue in N1–N3 are plausibly cited as justifying our beliefs in these constitutional necessities? It is common to separate metaphysical and epistemological questions, but the question I am asking has both metaphysical and epistemological aspects. What I am asking is: what *metaphysical* facts are plausibly cited as *justifying* our beliefs in N1–N3? More specifically: what metaphysical facts about the entities at issue in N1–N3 are, first, such that their holding would plausibly ground the truth of the associated claim, and second, such that we are plausibly in position to justifiably believe these facts to be in place?[14]

Call the sort of facts that would answer my question the 'justificatory' facts. I'll now argue that the non-Humean has metaphysically informative and epistemologically plausible accounts of the justificatory facts at issue in N1–N3, which accounts presuppose that there are causal necessities. By way of contrast, the Humean, in rejecting causal necessities, can only offer accounts of the justificatory facts at issue in N1–N3 that are metaphysically unilluminating and epistemologically implausible. The best accounts of what metaphysical facts are plausibly cited as justifying our beliefs in certain constitutional necessities thus presuppose that there are causal necessities. I will conclude that *Constitutional* → *Causal* is true: if one accepts any of the earlier constitutional necessities (or any of the multitude of variations on their themes), one should accept causal necessities.

2.1 The Non-Humean's Accounts

Let us start with the non-Humean's accounts of the justificatory facts at issue in N1–N3.

Necessarily, anything that is scarlet is red. Claims about the connections in paradigmatic cases of the determinable/determinate relation are typically thought to be justified *a priori*, given competence with the constitutive terms or concepts. This much doesn't settle the question of what justificatory facts are at issue in N1, however. As Williamson notes, to say that a truth is analytic or otherwise *a priori* itself leaves all the epistemological questions open:

> [Metaphysical accounts of analyticity, as truth in virtue of meaning, or in virtue of synonymy with a logical truth] provide no reason to regard

> analytic truths as in any way insubstantial. Even if core philosophical truths are analytic in such a sense, that does not explain how we can know or justifiably believe them. (Williamson 2007: 53)

Relatedly, to say that a truth is analytic or otherwise *a priori* leaves all the metaphysical questions open:

> [A]nalytic truths are not supposed to be always about words or concepts, even if words or concepts are supposed to play a special role in explaining their truth. The sentence 'Vixens are female foxes' is in no useful sense about the word 'vixen' or any other words; it is about vixens, if anything. (Williamson 2007: 48–9)

N1 is similarly not about words or concepts, but about certain natural properties, which we can ostend. So, granting that the truth of N1 may be established by attention to linguistic or conceptual phenomena, the questions remain: first, what metaphysical facts about the entities at issue in N1 are such that expressions for or concepts applying to these entities incorporate their necessary connection, and second, how might such facts be revealed in *a priori* deliberation, of whatever variety? An account of the justificatory facts at issue in N1 aims to answer both these questions.

The non-Humean's answer to the first question, concerning the metaphysical ground of N1, begins by registering three plausible claims, with which the Humean can agree. First, we have knowledge of the actual causal profiles of *being red* and *being scarlet*—that is, of what effects these properties, when instanced in certain circumstances, can enter into producing (notably, though of course not exclusively, as relevant to various of our experiences).[15] If we like, we can say that the actual causal profile of a property specifies the causal powers actually had or bestowed by the property, so long as such talk of powers is understood in metaphysically neutral fashion, as simply a way of registering the facts about actual causal potentialities. Second, we actually individuate these properties, as with most broadly scientific properties, by reference to their actual causal profiles: barring haecceitistic exceptions,[16] properties differing with respect to their causal profiles are different properties. Third, the causal profile of *being red* is actually contained in the causal profile of *being scarlet*: any effect that an instance of red can bring about in certain circumstances, in virtue of being red *simpliciter*, is an effect an instance of scarlet can bring about when in those circumstances, reflecting the fact that to be scarlet is to be red, in a specific way.[17] These claims support the claim that actually, anything that is scarlet is red. The non-Humean will additionally maintain, compatible with their denial of Hume's Dictum, that the actual causal profiles of *being red* and *being scarlet* are modally stable, such that these properties, when instanced in other worlds, have causal profiles that are the same as their actual profiles. As such, the causal profile of *being red* will be necessarily contained in the causal profile of *being scarlet*, supporting the claim that necessarily, anything that is scarlet is red.[18] Such facts about a necessary overlap in modally stable causal profiles provide a metaphysically straightforward and informative ground for the truth of N1.

The non-Humean answers the second question, concerning our epistemological access to the metaphysical ground of N1, as follows. As earlier, we have access to and actually individuate *being scarlet* and *being red*, as instanced throughout space and time, in terms of their causal profiles. Justified belief in N1, however, requires access to modal facts—in particular, facts about how these properties are individuated in modal contexts. How can *a priori* investigation reveal these individuation conditions, and the associated facts grounding N1? Here the non-Humean can appeal to the default assumption that our terms and concepts for broadly scientific properties incorporate the individuation conditions that we actually use, as applying not just throughout space and time, but also modally. This assumption, appropriately generalized to any broadly scientific entities, is key to the non-Humean's strategy for handling N1–N3, so let's set it off here:

> *The default assumption*: Other things being equal, our terms and concepts for broadly scientific entities incorporate the conditions we actually use to individuate these entities, as applying not just throughout space and time, but also modally.

This assumption is reasonably considered the default, on grounds of being the simplest and most straightforward extension of our actual individuation conditions for broadly scientific properties to modal contexts.[19] Of course, here as per usual, *The default assumption* might be overturned, given good reason; hence the caveat 'other things being equal'. I will later argue that there isn't any good reason on the Humean's table. For the present my point is simply that, given *The default assumption* and the fact that we have knowledge of the actual individuation conditions of *being scarlet* and *being red*, the non-Humean has an epistemologically plausible story to tell about how *a priori* investigation into the terms or concepts for these properties can result in justified belief in N1.

Necessarily, anything having a certain mean molecular kinetic energy (MMKE) has a certain temperature. N2 may differ from N1 in being an *a posteriori* necessity such that, given that MMKE is actually identical to or actually realizes temperature, MMKE is necessarily identical with or necessarily realizes temperature.[20] Alternatively, one might suppose that the necessity at issue is purely *a priori*, with empirical investigation entering only into the 'concept-formation' stage. Even supposing that N2 is supposed to be, or to follow from, an *a posteriori* necessity, however, it will still be the case that (as per the usual understanding of the epistemology of such necessities) empirical investigation is required mainly to establish that the entities *actually* stand in the relation at issue; so far as the modal aspect of the claim is concerned, this is still a matter of broadly *a priori* deliberation. So, for example, while empirical investigation is needed to determine that (as the case may be) MMKE is identical with or realizes temperature, it is *a priori* that *given* that MMKE is identical with or realizes temperature, this is necessarily the case. Such an epistemological stance on *a posteriori* necessities reflects what, post-Kripke, is a minimal departure from the traditional empiricist view that what is necessary is *a priori*. Indeed, this stance was endorsed by Kripke himself: on his preferred account, the modal aspect of *a posteriori* necessities (whether expressing identities or natural kind essences) is established by *a priori* consideration of how concepts or terms for the relevant entities are applied in hypothetical scenarios (not, note, by relying on the broadly logical proof of the necessity of identity—though that would be another *a priori* route to the modal aspect of identity claims).

Given all this, what is the non-Humean's answer to the first question, concerning the metaphysical ground of N2? Their answer will be along lines similar to that given for N1, starting with three claims that the Humean can accept. First, the non-Humean will claim that we associate the properties (or states) *having a certain MMKE* and *having a certain temperature* with certain actual causal profiles. Second, as earlier, the non-Humean will claim that (barring haecceitistic exceptions) we actually individuate broadly scientific properties (state types) via these actual causal profiles. Third, the non-Humean will claim that we have knowledge of the fact that the causal profile of *having a certain MMKE* is identical with or contained in the causal profile of *having a certain temperature*. The previous claims support the claim that actually, anything that has an MMKE has a temperature. The non-Humean will additionally maintain, compatible with their denial of Hume's Dictum, that the causal profiles of *having a certain MMKE* and *having a certain temperature* are modally stable, such that these properties (state types), when instanced in other worlds, have causal profiles that are the same as the properties actually have. As such, the causal profile of *having a certain temperature* will be necessarily identical with or contained in that of *having a certain MMKE*, supporting the claim that necessarily, anything that has an MMKE has a temperature. Such facts about a necessary overlap in modally stable causal profiles provide a straightforward and informative metaphysical ground for the truth of N2.

The non-Humean's answer to the second question, concerning our epistemological access to the metaphysical ground of N2, also proceeds as in the case of N1; that is, by reference to *The default assumption*, according to which other things being equal, our terms and concepts for broadly scientific property or state types incorporate modal individuation conditions that are the same as those we use to actually individuate these types throughout space and time. Given this assumption, and the fact that we have knowledge of the actual individuation conditions of *having* *an MMKE* and *having a temperature*, the non-Humean has a plausible epistemological story to tell about our access to the metaphysical facts grounding N2.

Necessarily, electrons are negatively charged. Like N2, N3 might be seen as an *a posteriori* necessity; alternatively, perhaps N3 is *a priori*, and the empirical facts enter simply into forming the concepts at issue. Either way, the non-Humean's account of the justificatory facts proceeds along the same lines as for N1 and N2. First, the non-Humean will claim that we have knowledge of the causal profiles of the particular type *electron* and the property type *being negatively charged*. (Though causal powers and profiles are usually associated with properties, we can and do also associate powers and profiles with particular types. Claim three expresses one straightforward way in which the causal profiles of particular and property types may be related.) Second, they will claim that we actually individuate tokens of the particular type *electron* by reference to the associated causal profile, and actually individuate tokens of the property type *being negatively charged* in terms of the associated actual causal profile. Third, they will claim that we come to learn, as a matter of empirical fact, that the causal profile associated with *being negatively charged* is contained in the causal profile associated with *electron*.[21] These claims in turn support the claim that actually, every electron is negatively charged. The non-Humeans will again maintain, compatible with their denial of Hume's Dictum, that these causal profiles are modally stable, providing support for the claim that necessarily, every electron is negatively charged. And once again, the non-Humean can tell a plausible epistemological story about how this modal stability is revealed in *a priori* deliberation, by appeal to *The default assumption*, according to which our modal principles of individuation incorporate the actual such principles to which we uncontroversially have access.

2.2 The Humean's Accounts

Next, let's turn to what accounts the Humean can give of the justificatory facts at issue in N1–N3. I'll engage in some comparative assessment along the way.

Necessarily, anything that is scarlet is red. Let's start by asking: can the Humean implement the non-Humean's strategy of appeal to a modally stable overlap in causal profiles of the properties involved? On the face of it, no. After all, according to Hume's Dictum, there are no causal necessities; hence the causal profiles of properties are not modally stable; hence notwithstanding that the causal profiles of these properties actually overlap (as Humeans will agree), the Humean has no reason to suppose that they necessarily do so.

The Humean might nonetheless attempt to locate the necessity of N1 in a necessary overlap of causal profiles, in two ways.

First, the Humean might suggest that the necessity at issue in N1 is grounded in its being the case that *whatever* causal profile *being scarlet* happens to have at a world, this causal profile would contain that of *being red*. The suggestion is, however, a non-starter: given that, as per Hume's Dictum, there are no modal restrictions on the causal profiles of either *being scarlet* or *being red*, there is no reason to think that the causal profile of *being red* will necessarily be contained in the causal profile of *being scarlet*. (I won't bother revisiting this strategy when discussing N2–N3.)

A second, somewhat more principled way of implementing the 'overlap' strategy would be to maintain that colour properties are not subject to Hume's Dictum, properly understood. Recall that Humeans might respond to Bird by restricting the application of Hume's Dictum to sparse (natural, fundamental) properties—allowing, in particular, that non-fundamental (e.g. functional or structural) properties can have stable modal profiles, insofar as what it is to be (an instance of) such a property is effectively to be capable of filling a certain causal role. If Hume's Dictum is so restricted, and if colour properties are non-fundamental (e.g. functional/structural), then the Humean could maintain that the properties *being scarlet* and *being red* have modally stable, necessarily overlapping causal profiles.

But colours are not obviously non-fundamental structural properties (see, e.g., Campbell 1993; Yablo 1995; Watkins 2002). To be sure, many Humeans are also physicalists, who will suppose that colour properties (at least, understood as appearance properties that are at least partly psychological) are not fundamental. Pending the outcome of debate on the nature of colours, however, there is no guarantee that the Humean can tell the sort of metaphysically informative, epistemologically plausible story that the non-Humean can tell about the justificatory facts concerning N1. And in any case, the non-Humean's story will retain the advantage that it does not antecedently require commitment on whether colours are fundamental (or, relatedly, to physicalism).

A different metaphysical account of the necessity at issue in N1, having nothing to do with overlapping causal profiles, is available to the Humean. Here the account of the justificatory facts appeals to *quiddities*—primitive identities that are the property equivalent of haecceities—of the properties at issue.[22] The suggestion applied to N1 would be that *being scarlet* and *being red* have certain quiddities, which are, as it happens, necessarily connected. Quiddities float free of causal profiles, just as haecceities do; hence the posit of necessarily overlapping or otherwise connected quiddities would provide a metaphysical ground for the necessity of N1, even granting, as per Hume's Dictum, the modal instability of the properties involved. And presumably properties may have quiddities, whether or not they are fundamental.

Indeed, proponents of Hume's Dictum often suppose that properties have quiddities (see Armstrong 1989: 44; Schaffer 2004; and Lewis 2009)—as needed, presumably, to ground trans-world identity of properties in the absence of a stable causal profile. But there are problems with a quiddity-based account of the justificatory facts at issue in N1, from both a metaphysical and epistemological point of view. An account of the necessary connection at issue in N1 in terms of necessarily connected quiddities is metaphysically unilluminating, since there is nothing in the nature of quiddities (being primitive property identities) which indicates that, much less illuminates why, *being scarlet* and *being red* are necessarily connected.

A quiddity-based account is also epistemologically implausible. What we seek in an account of the justificatory facts of a given claim is not only metaphysical illumination, but also a plausible story about our epistemological access to the relevant metaphysical facts. But there is not, so far as I can tell, any plausible story to be told about our access to facts involving quiddities, much less to facts about overlapping quiddities of the sort that might ground constitutionally necessary connections. We don't actually perceive quiddities (perception being a causal affair), nor can I see how we might have access to such entities via conceivability, rational insight, or any form of modal perception or intuition. Nor will it help if we allow (as I think we should) that *a priori* deliberation may incorporate principles of abductive explanation, with the outcome of such deliberation being something like a theoretical inference to the best explanation; for there doesn't appear to be any theoretical motivation for quiddities. Science provides no such motivation: to the extent that the relevant terms receive scientific definitions, these are exclusively in terms of their actual causal profiles.[23] And, at least at present, there isn't any philosophical motivation for incorporating non-causal quiddities into the expressions or concepts for the entities at issue—at least, no philosophical motivation independent of antecedent commitment to Hume's Dictum, the appropriateness of which we are presently investigating. The best attempt at independent motivation for quiddities, due to Armstrong, has the form of an argument by analogy:

> [S]woyer . . . argues that properties must have 'essential features' [namely] the relations of 'nomic implication' which properties have to other properties. But why need properties have essential features at all? Perhaps their identity is primitive. To uphold this view is to reject the Principle of the Identity of Indiscernibles with respect to properties. Properties can be different, in the same way that, many of us would maintain, ordinary particulars can just be different although having all their features in common . . . properties can be their own essence. (Armstrong 1983: 160)

More to the present point, to allow that properties have a primitive identity is to reject the Distinctness of Discernibles: properties can be the same in spite of having completely different causal profiles, Armstrong suggests, just as ordinary particulars can be the same in spite of having completely different properties. But the analogy, hence the argument, fails (see Wilson

2005). There is an 'inference to the best explanation' case for thinking that some broadly scientific particulars have haecceitistic natures; in particular, we have actual experience of persons persisting through relatively extreme changes in their properties, as when a single human moves from infancy to adulthood—experience of which the thesis that particulars have haecceities is, perhaps, the best explanation. But we do not, in either ordinary or scientific contexts, experience or posit properties as persisting through any but very minor changes in their actual causal profiles (haecceitistic exceptions, such as *being Barack Obama*, aside). There is no parallel motivation for thinking that properties have an identity independent of their causal profile; hence there is nothing to explain, such that the posit of quiddities would be the best, or at any rate a reasonable, explanation. So the analogy fails, and with it what independent philosophical motivation exists for postulating non-causal quiddities, and for taking the necessity at issue in N1 to be grounded in such quiddities.

This exhausts the Humean's available options, it seems, so far as providing an account of the justificatory facts at issue in N1 is concerned. There are two: ground N1 in a necessary overlap in the causal profiles associated with *being red* and *being scarlet*, or ground N1 in a necessary overlap in (non-causal) quiddities associated with these properties. The first strategy (also endorsed by the non-Humean) would provide a metaphysically informative and epistemologically plausible account of the justificatory facts, but at the cost (which the non-Humean does not have to pay) of commitment to the controversial view that colours are structural, non-fundamental properties. The second is metaphysically unilluminating (*why* do the primitive identities of *being scarlet* and *being red* overlap?) and epistemologically implausible. I conclude that the non-Humean has a better account of the justificatory facts at issue in N1 than the Humean.

Variations on the preceding themes apply to the cases of N2 and N3—only in these cases the Humean's options are yet more limited.

Necessarily, anything with a certain mean molecular kinetic energy (MMKE) has a certain temperature. Again, *prima facie*, the Humean is not in position to implement the non-Humean's strategy of grounding the necessity at issue in N2 in a necessary overlap in causal profiles. Moreover, here there is no hope of achieving such a necessary overlap via commitment to the properties at issue's being non-fundamental structural properties, for *having a certain MMKE* is not appropriately seen as a structural property. Rather, it is a mere mathematical average of an additive intrinsic property (kinetic energy) of whatever fundamental entities compose molecules. As such, even the 'sparse' Humean is committed to *having a certain MMKE*—and indeed, to all properties that are additive functions of intrinsic properties of fundamental entities, such as *having a certain mass*—being such as to have modally unstable causal profiles.

As such, it appears that the only available Humean account of the justificatory facts at issue in N2 is one appealing to a necessary overlap in quiddities, which option is, as earlier, metaphysically unilluminating and epistemologically implausible. Again, the non-Humean does better.

Necessarily, electrons are negatively charged. Again, the Humean's endorsement of Hume's Dictum appears to prevent their giving an account of the necessity at issue in terms of a necessary overlap in causal profiles. And as with N2, the appearance is genuine, for there is no hope here of implementing the overlap strategy by appeal to the non-fundamentality of the entities involved, for *being an electron* and *being negatively charged* are fundamental broadly scientific entities—the sort of entities to which Hume's Dictum is supposed to apply (that is, the sort of entities whose causal profiles may modally vary), if any are.

Hence it seems that the Humean can ground N3 only by appeal to a purported quiddity associated with *being negatively charged* standing in some necessary connection to (haecceities of?) electrons, which account is again both metaphysically unilluminating and epistemologically implausible. Again, the non-Humean does better.

3. FROM CONSTITUTIONAL NECESSITIES TO CAUSAL NECESSITIES

Let's recap and draw some conclusions. The most promising case to be made in support of the Humean's differential treatment of constitutional and causal necessities is one appealing to *Constitutional necessity*, according to which the only necessities are constitutional necessities. But a closer look at what facts might plausibly justify certain constitutional necessities involving broadly scientific entities indicates that *Constitutional necessity* is false. In particular, in each of N1–N3, the best account of what facts about the entities involved are plausibly cited as justifying the constitutional necessity at issue presupposes, contrary to Hume's Dictum, that entities involved have distinctive, modally stable causal profiles—that is, that there are necessary causal connections between wholly distinct, intrinsically characterized entities.

These considerations indicate that the following conditional is true:

> *Constitutional* → *Causal*: If one accepts certain constitutional necessities, one should accept certain causal necessities.

What is the bearing of this result on a comparative assessment of the Humean and non-Humean frameworks in the metaphysics of science?

To start, it is worth noting that the considerations supporting *Constitutional* → *Causal* are relatively independent of the dispute between Humeans and non-Humeans. The overall dialectic is as follows. There are certain constitutional necessities, typified by N1–N3, holding between broadly scientific, not-wholly-distinct entities, that are very commonly accepted, and that in any case Humeans are not in the business of denying and typically suppose they can accept. Humeans and non-Humeans are then invited, so to speak, to each provide their best account of the justificatory facts supporting these constitutional necessities. Independent considerations indicate that the non-Humean's accounts of these facts, each appealing to a necessary overlap in causal profiles of the entities involved, are in each case metaphysically informative and epistemologically plausible. By way of contrast, the Humean's best account of N1 (appealing to a necessary overlap in causal profiles) is, while metaphysically informative and epistemologically plausible, purchased at the price of commitment to a controversial account of colours—a price the non-Humean need not pay; and the Humean's best accounts of N2 and N3 (appealing to quiddities) are both metaphysically unilluminating and epistemologically implausible. In each case, the non-Humean's account fares better. We should accept the holding of those metaphysical facts that enter into the best account of the justificatory facts concerning claims we accept. Hence if we accept certain constitutional necessities (namely, any of the sort typified by any of N1–N3), we should accept certain causal necessities.

Given that the constitutional necessities at issue involve broadly scientific entities, this result seems to favour non-Humeanism over Humeanism, so far as the metaphysics of science is concerned. We can more firmly establish this bearing by noting that the Humean has only two options for responding to the result, each of which is quite unattractive.

The first is to deny *Constitutional* → *Causal*, maintaining acceptance of constitutional necessities such as N1–N3 by insisting on retaining the Humean's less-than-satisfactory accounts of the justificatory facts. Such a response obviously comes at a cost in plausibility of the Humean's view. By way of contrast, the non-Humean's endorsement of N1–N3 comes at no such cost.

The second response is for the Humean to reject constitutional necessities such as N1–N3 (which again are ubiquitously multiplied). Any such denial also comes at a cost for the Humean, since constitutional necessities along lines of N1–N3 are highly plausible and commonly accepted—indeed, are typically rejected only by those endorsing various kinds of scepticism about analyticity, *a posteriori* necessities, or essential properties of even the benign sort associated with natural kinds. Relatedly, as noted, proponents of Hume's Dictum have typically supposed that its application to causal necessities doesn't require rejecting constitutional necessities as regards to such entities. To the extent that Humeans respond to *Constitutional* → *Causal* by rejecting any constitutional necessities undermining their endorsement of Hume's Dictum, their position seems not only extensionally incorrect but *ad hoc*. By way of contrast, the non-Humean is under no pressure to reject any of these constitutional necessities, and moreover their acceptance is naturally situated in their overall metaphysics.

The Humeans fares worse than the non-Humeans, no matter how they respond. I conclude, then, that *Constitutional* → *Causal* provides needed leverage in a comparative assessment of Humean vs. non-Humean frameworks in the metaphysics of science, in the non-Humean's favour[24].

NOTES

1. Somewhat more precisely, Hume's Dictum states that there are no metaphysically necessary connections between wholly distinct, intrinsically characterized entities. See Wilson forthcoming for further precisifications and interpretive options.
2. For this to make sense, the background condition may include states of affairs (e.g. the absence of countervailing forces and the continued existence of the world) but may not include the actual laws of nature.
3. Or so the story goes. I hedge here since it seems to me a live possibility that in the first instance, the term 'water' is intended (e.g. by hydrologists and ordinary language users) to mark a broadly functional, higher level kind, such that 'being H_2O' isn't in fact part of the meaning of 'water'. That said, I'm happy to allow that higher level natural kind terms may incorporate lower level aspects of actual reality as part of their intended meaning, and here grant for the sake of argument that water is such a term.
4. Hence it is that Schaffer does not find it necessary to consider any of the broadly metaphysical and epistemological considerations that Shoemaker raises in support of his causal view of properties.
5. One might wonder if such constitutional necessities might be understood as cases of partial identity, with the proof of the necessity of the connection in question grounded, somehow or other, in that available for the case of identity. But it is unclear how the indirect proof would go, and in any case accounts of the preceding relations in terms of partial identity are in short supply (with Baxter's 1988 and 2001 accounts of parthood and instantiation and parthood being notable exceptions) and generally seen as unacceptably revisionary (requiring, e.g., rejection of Leibniz's law). Some suggest that parthood is analogous to partial identity (see Lewis 1991; Sider 2007); but it is even less clear how the necessity of a relation that is merely analogous to (partial) identity can be grounded in that of identity. So there is little hope of gaining the necessity of the preceding constitutional relations by forcing them into the mold of identity. Concerns about "compelling proof" aside, similar remarks apply to the mereological principles floated in Cameron (forthcoming), as providing Humean-friendly reason to accept necessary connections between 'overlapping' entities. In fact, Cameron doesn't think such principles—most promisingly, compositional essentialism—are true, since on his very general understanding of parthood, wholes and parts may come apart (e.g. a cup might exist through a change in its composing molecules). I don't think these cases bear on the parthood relation targeted by classical mereology, but even supposing compositional essentialism were generally true there is little reason to think that constitutional relations can generally be forced into the mold of mereology.
6. Again, so the story goes. If logic is permeable to empirical or other considerations (as per Quine 1951/53), then even a broadly logical proof may fail to be absolutely compelling.
7. For this reason, while I agree with Cameron forthcoming that necessary connections acceptable to the Humean must be grounded in ontological

dependence, I disagree that the dependence in question should be further cashed in terms of truthmakers.

8. How might we be justified in believing that entities are constitutionally connected, when they are? I won't here provide a general answer, on the Humean's behalf or otherwise. Shortly, however, I will consider what facts about the entities at issue in certain commonly accepted constitutional necessities might plausibly justify our beliefs in these necessities.
9. Even dispositional essentialists will typically agree that causes and effects do not constitute each other in the specified sense, since, first, they will typically agree that causes and effects are diachronic, and second (and more to the metaphysical point) they will typically agree that a token of a given cause-type may exist without a token of its usual effect-type's existing (if requisite background conditions are not in place), and that a token of a given effect-type may exist without a token of its usual cause-type's existing (if caused by a token of some other cause-type).
10. One may wonder if *being negatively charged* is an intrinsic property. Answer: yes, on all standard accounts of 'intrinsic'. So, for example, an entity could be negatively charged even if it was a 'lonely' entity (the only entity existing at a world). More specifically, the having of this property by an entity does not entail that the entity stands in any causal relation *R* to another entity (at a minimum—or so Humeans and non-Humeans can agree—due to the failure to hold of requisite background conditions). Hence Lewis (1986: 77) says, '[w]e tend to think that positive and negative charge are perfectly natural intrinsic properties of particles'.
11. Alternatively, N2 might be rejected on grounds that temperature is strongly emergent from, rather than identical to or realized in, MMKE. A similar option is available to those denying that mental states are physically realized; however, it seems reasonable to assume that temperature is not strongly emergent.
12. On the subject of anti-essentialism: one may wonder if Lewis, a paradigmatic contemporary Humean, is committed to anti-essentialism as a consequence of his counterpart-theoretic account of the truth of *de re* modal claims. Answer: no, not in any sense relevant to the present discussion. On Lewis's counterpart theory, there are appropriate contexts in which N3 (as well as N1 and N2) is true, and the discussion to follow may be seen as restricted to such contexts.
13. Among other things, the latter, stronger claim commits one to denying that there are any fundamental natural kinds, given that membership in any fundamental natural kind requires, at least in part, the essential having of certain intrinsic properties.
14. This question presupposes a broadly realist semantics and ontology for N1–N3. This assumption begs no questions here, since Humean and non-Humean parties to the present debate are all realists, and indeed, broadly scientific realists. See, e.g., Lewis 1970 and Bird 2007.
15. Of course, at any given point of inquiry we may not be in completely accurate possession of these actual causal profiles, but what follows won't turn on such issues, nor on the further complex and broadly scientific matter of how causal profiles are assigned to broadly scientific entities (similarly when considering N2 and N3). Note also that the claim that colour properties have actual causal profiles doesn't entail anything about whether colours are functionally characterizable in non-qualitative terms; perhaps the production of certain qualitative experiences is an irreducible part of the causal profiles at issue.
16. Such as *being Barack Obama*, a case I'll discuss further down the line.

17. Moreover, there is plausibly a proper containment (focusing on individual powers: proper subset) relation between the power profiles of *being red* and *being scarlet*, reflecting that in virtue of being more determinate, instances of *being scarlet* can do some things that instances of *being red* can't do. Hence it is, as per the usual understanding of the determinable/determinate relation, that the entailment in the direction from *being red* to *being scarlet* does not go through.
18. This claim is compatible with some degree of variation in causal profile, so long as it is principled and contained.
19. Additional motivation for *The default assumption* might advert to the parallels between claims about future possibilities and other modal claims.
20. Constitutional realization claims, if understood as *a posteriori* necessities, might be best seen as analogous to essence claims of the sort at issue in N3, such that if it turns out that MMKE actually realizes temperature, it necessarily does so.
21. This claim registers the general fact that causal profiles of particular types can be factored into or distributed over the causal profiles of property types.
22. Here and throughout, I assume that the quiddities to which the Humean may appeal are non-causal, in not being associated with specific causal powers or profiles as any part of their nature.
23. Again, allowing for some principled and contained variation.
24. Thanks to Helen Beebee, Benj Hellie, Nigel Leary, Francis Longworth, Jonathan Schaffer, and participants of the 2007 AHRC Metaphysics of Science conference 'Nature and its Classification', for helpful comments.

REFERENCES

Armstrong, D. M. (1983) *What Is a Law of Nature?* Cambridge: Cambridge University Press.

———. (1989) *A Combinatorial Theory of Possibility*. Cambridge: Cambridge University Press.

Baxter, D. L. M. (1988) 'Many-one identity', *Philosophical Papers*, 17: 193–216.

———. (2001) 'Instantiation as partial identity', *Australasian Journal of Philosophy*, 79: 449–64.

Beebee, H. (2002) 'Contingent laws rule', *Analysis*, 62: 252–5.

Bird, A. (2001) 'Necessarily, salt dissolves in water', *Analysis*, 61: 267–74.

———. (2007) *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.

Cameron, R. (forthcoming) 'From Humean truthmaker theory to priority monism', *Noûs*.

Campbell, J. (1993) 'A simple view of colour', in J. Haldane and C. Wright (eds), *Reality: Representation and Projection*, Oxford: Oxford University Press, 257–68.

Kripke, Saul (1972/1980) *Naming and Necessity*. Cambridge: Harvard University Press.

Lewis, D. K. (1970) 'How to define theoretical terms', *The Journal of Philosophy*, 67: 427–46.

———. (1991) *Parts of Classes*. Oxford: Oxford University Press.

———. (2009) 'Ramseyan humility', in D. Braddon-Mitchell and R. Nola (eds), *Conceptual Analysis and Philosophical Naturalism*, Cambridge, MA: MIT Press, 203–22.

Lowe, E. J. (2006) *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford: Oxford University Press.

Quine, Willard Van Orman (1951/53) 'Main trends in recent philosophy: two dogmas of empiricism', *The Philosophical Review*, 60: 20–43. Reprinted, with alterations, in Quine 1961.

———. (1961) *From a Logical Point of View*, 2nd ed. Cambridge, MA: Harvard University Press.

Schaffer, J. (2004) 'Quiddistic knowledge', in F. Jackson and G. Priest (eds), *Lewisian Themes*, Oxford: Oxford University Press 210–30.

Shoemaker, S. (1998) 'Causal and metaphysical necessity', *Pacific Philosophical Quarterly*, 79: 59–77.

Sider, T. (2007) 'Parthood', *Philosophical Review*, 116: 51–91.

Watkins, M. (2002) *Rediscovering Colors: An Essay in Pollyanna Realism*. Dordrecht, Netherlands: Springer.

Williamson, T. (2007) *The Philosophy of Philosophy*. Oxford: Blackwell.

Wilson, J. (2005) 'Supervenience-based formulations of physicalism', *Noûs*, 39: 426–59.

———. (forthcoming) 'What is Hume's Dictum, and why believe it?' *Philosophy and Phenomenological Research*.

Yablo, S. (1995) 'Singling out properties'. *Philosophical Perspectives*, 9: 477–502.

12 Realism, Natural Kinds, and Philosophical Methods

Richard N. Boyd

1. INTRODUCTION: TOWARDS A FULLY DEVELOPED PHILOSOPHICAL NATURALISM

Many aspects of recent realist work in the philosophy of science have some affinities with philosophical naturalism. Realists emphasize that the methods of science are theory-dependent, so that the justification for those methods in any particular application will have an *a posteriori* scientific component. Realist treatments of natural kinds treat the definitions of those kinds as at least partly *a posteriori*, and 'causal' or 'naturalistic' theories of reference for natural kind terms similarly treat at least some aspects of reference as matters for *a posteriori* empirical investigation. In a number of papers I've been developing arguments to the effect that, properly developed, scientific realism dictates a thoroughgoing, anti-foundationalist, anti-reductionist naturalistic approach to philosophical matters. What I propose to do here is to pull together the core of those arguments and to indicate what sort of philosophical naturalism is implied.

I'll begin by laying out what I take to be the appropriate realist treatments of theory-dependent methods and of natural kinds and naturalistic theories of reference, sketching the arguments for them and citing papers which offer more sustained arguments. I'll then concentrate on exploring the philosophical and metaphilosophical implications of the realist positions I've laid out. Of course, readers with different conceptions of realism may have different views about its implications, and some readers initially sympathetic to my conception of realism may come to reject it in light of its implications. Still, I hope that the considerations I rehearse here will lead to productive discussions.

2. REALIST PHILOSOPHY OF SCIENCE

2.1 Projectibility and Evidence

It is widely recognized by philosophers and psychologists who may differ sharply about other matters that when scientists (and ordinary people) address a scientific question they seek answers from a "small handful" of options commended to them by the paradigm or framework, or tradition within which they operate (see, e.g., Fine 1984; Kuhn 1970; Goodman 1973; Koslowski 1996; Lipton 1991, 1993; Quine 1969). These are the projectible answers to that question (in the sense of Goodman 1973).

Projectibility judgments are judgments of scientific plausibility informed by background theories and (partly tacit) beliefs and concepts and obtained by partly (sometimes largely) tacit inferences characteristic of the inferential architecture of the relevant discipline(s): by the practices licensed by the relevant paradigm (in the sense of Kuhn 1970). They play an absolutely fundamental role in the evaluation of evidence for or against proposed theories. To an extremely good first approximation a theory, *T*, counts as confirmed by evidence, *E*, given the evidential standards of a scientific community at a time just in case:

1. *T* is projectible by that community's standards, and
2. *E* favors *T* over all relevant alternatives (that is, rival theories also projectible by those standards), and
3. *E* was gathered so as to control for experimental or observational artifacts suggested by theories projectible by those standards.

2.2 Projectibility and Truth

What has projectibility to do with truth? Here's another way of putting these standards evaluating evidence:

> *Basic methodological rule of science:* Carefully choose from among relevant alternatives (that is, theories recommended by best current theories), controlling for effects suggested by best current theories.

Why is this rule reliable? In particular, what makes relying on current theories and inferential practices in this way reliable? Answer: Surely *not* their currency. Instead, to see why (and when) these methods are reliable, we need yet more rules.

> *Basic 'falsificationist' rule for objective testing:* Try to falsify *T*. Test *T* where it's most vulnerable, i.e. under circumstances where it's most likely to go wrong if it's not (relevantly, approximately) true.
>
> *Rule for identifying such circumstances:* Identify the most plausible alternatives to *T*: its projectible rivals. Similarly for identifying experimental artifacts.

So in practice scientists rely on projectibility judgments in order to subject theories to rigorous testing. If they don't at first, then, pretty often, colleagues or journal referees insist that they do. So when is this sort of scientific practice epistemically reliable? To a pretty good first approximation (better ones later) the basic methodological rule is reliable just when the prevailing theories are accurate enough and the prevailing inferential practices are reliable enough that (a) pretty often an approximately correct answer to a scientific question will be among the projectible answers to it, and (b) pretty often one can avoid experimental artifacts by controlling for those suggested by projectible theories.

The basic idea is that theory-dependent methods work just in case the relevant background theories and methods are accurate and reliable (see Boyd 1983, 1985a, 1985b, 1989).

2.3 Natural Kinds and Reference

The Basic *Accommodationis*t Picture.

Paul Sherman and his associates (Sherman 1977; Sherman and Reeve 1997) have identified two different sorts of alarm call in Belding's ground squirrels. One sort of call warns of an aerial predator; the other warns of a terrestrial one. The squirrels respond differently to the two sorts of call, engaging in different evasive behaviours appropriate to the different sorts of predators. Sherman and Reeve have thus confirmed a semantic hypothesis about those calls. One, call it *a*, refers to and warns about aerial predators; the other, *t*, refers to and warns about terrestrial predators.

What underwrites the referential semantics of *a* and *t*? It's by no means the case that *a*s are invariably associated with aerial predators or *t*s with terrestrial ones; there are false positives and false negatives. Instead ,what makes *a* and *t* refer to aerial and terrestrial predators, respectively, is instead that (1) there is a tendency for *a*s to be produced in response to aerial predators rather than to other survival-relevant features of the squirrels' environment, (2) there is a similar tendency for *t*s to be produced in response to terrestrial predators, and (3) these facts figure centrally in explaining how Belding's ground squirrels avoid predation.

In this case the referential hypothesis functions as a component in the explanation of an *achievement*—predator avoidance—on the part of the relevant organisms. It helps to explain how the perceptual and cognitive structures in Belding's ground squirrels are *accommodated* to relevant causal features of their environment so as to facilitate predator avoidance. What I have proposed (See Boyd 1989, 1990, 1991, 1999a, 1999b, 2001b) is that the theory of reference for natural kind terms in science (and in everyday life) is likewise a component in the explanation of our epistemic achievements—our successes, such as they are, in induction and explanation. On this view the kind *natural kind* is itself a natural kind in the study of the epistemic reliability of human inductive and explanatory practices. What is to be explained is the ways in which the *accommodation* of classificatory and linguistic practices to causal factors in the world contributes to the reliability of those practices.

The Accommodationist Theory: Initial Approximation.

The fundamental question which the theory of natural kinds addresses is this: 'How do classificatory practices and their linguistic manifestations help to underwrite the reliability of scientific (and everyday) inductive/explanatory practices?' When we inquire about the definition of a natural kind, *K*, we're asking something like this: what commonalities in the causal profiles of things we classify as *K*s explain such inductive and explanatory successes as we have achieved? H_2O is the definition of the kind water because (1) to a good enough first approximation we tend to classify substances under the term 'water' (or related term in other languages) just in case they're mainly H_2O and (2) this fact helps to explain our inductive/explanatory successes with respect to the term 'water'.

Of course the definition of a natural kind, *K*, depends on the actual inferential practices of the relevant scientific communities: on the *inferential architecture* of the relevant discipline. So the definition of any given *K* depends on the characteristic inferential connections between the term referring to *K* and all of the other natural kind terms within the discipline. The correct referential semantics for discourse within a discipline will, to a good first approximation, be an assignment, to each natural kind term, of a family of properties such that (1) the actual usage of each term approximately 'tracks' the family assigned to it, and (2) the fact that this pattern of tracking occurs explains the reliability—such as it is—of the discipline's inferential practices. Here's a more precise way of saying all this.

Let *M* be a disciplinary matrix and let $t_1, \ldots, t_n$ be the natural kind terms deployed within the discourse central to the inductive/explanatory successes of *M*. Then the families $F_1, \ldots, F_n$ of properties provide definitions of the kinds referred to by $t_1, \ldots, t_n$, and determine their extensions, just in case:

1. (Epistemic access condition) There is a systematic, causally sustained, tendency—established by the causal relations between practices in *M* and causal structures in the world—for what is predicated of t_i within the practice of *M* to be approximately true of things which satisfy F_i, $i = 1, \ldots, n$. In particular, there is a systematic tendency for things of which t_i is predicated to have (some or most of) the properties in F_i.1
2. (Accommodation condition) This fact, together with the causal powers of things satisfying these explanatory definitions, causally explains how the use of $t_1, \ldots, t_n$ in *M* contributes to accommodation of the inferential practices of *M* to relevant causal structures. It explains whatever tendency there is for participants in *M* to identify causally sustained generalizations, to obtain correct explanations, or to obtain successful solutions to practical problems.

It is, of course, central to the accommodationist conception that the proper definition of a natural kind is an *a posteriori* question and that sound methodological practices require that scientists revise their conceptions of natural kinds in the light of new findings. The achievement of accommodation requires that we treat our conceptions of natural kinds as revisable (see Boyd 1990, 1991, 1999a, 1999b, 2001b; see also Hendry 2005).

Partial Denotation and Other Complications.

One respect in which the basic accommodationist conception of natural kinds and natural kind terms is only an approximation is, of course, that there will sometimes be natural kind terms that fail altogether to refer. More interesting are the phenomena of partial denotation and denotational refinement in the sense of Field 1973. Roughly, a term t partially denotes different kinds k_1 and k_2 in a disciplinary matrix M when the epistemic connection between the uses of t in M and k_1 explains very nearly the same achievements in M as does the connection between t and k_2. In practice, practitioners in M do not distinguish between k_1 and k_2; their use of t corresponds, in a sense, to something like the union of k_1 and k_2. Nevertheless, the reliability of their practice is compromised by this feature of their conceptual and linguistic practices. An improvement in reliability could be achieved by drawing the k_1–k_2 distinction and by replacing the existing use of t with the use of two terms (one of which might, but need not, be t), one referring to $k_{1,}$ and the other to k_2. This is *denotational refinement* in Field's sense. (For further discussions see Boyd 1999b, 2003a, 2003b.)

An obvious example involves the use of the term 'element' in chemistry before the distinction we now mark by the terms 'element' and 'isotope' was drawn. Another plausible example involves the term 'species' as it is used in biology. A number of biologists and philosophers have argued for 'pluralism' about the species category: the thesis that for different branches of biological inquiry—ecology, animal behaviour, evolutionary biology, etc.—different notions of species are required, but that this need is unrecognized in practice, so that biologists work with a not-fully-adequate conception.2 If this is true, then the term 'species' partially denotes each of several kinds of biological kinds and denotational refinement is in order.

Homeostatic Property Cluster (Hpc) Natural Kinds.

Here's another complication. We ordinarily think of the extension of a natural kind term as a set of things. Sometimes, however, the satisfaction of the accommodation condition requires that the a natural kind be defined by a naturally occurring clustering of properties with the consequence that (1) it lacks precisely defined membership conditions and, sometimes (2) the properties in the defining cluster vary over time and/or space. The resulting 'vagueness' in the extension of the associated kind term reflects not an inappropriate imprecision but a precise accommodation of classificatory practices to relevant causal phenomena. Biological species are paradigmatic HPC natural kinds.

It follows from evolutionary theory that they will ordinarily lack completely determinate boundaries, so any precisification of a definition of a species would misrepresent biological reality and thus undermine accommodation. (For further discussions see Boyd 1999a, 1999b, 2003a, 2003b.)

A Kind of Relativism.

It follows from the accommodation theory that the naturalness of a natural kind is discipline relative. There are not kinds which are natural simpliciter but instead kinds that are natural with respect to the inferential architectures of particular disciplinary matrices. Any talk of natural kinds, properly understood, involves (perhaps tacit) reference to or quantification over disciplinary matrices (Boyd 1999b, 2001b).

Accommodation, Broadly Understood.

We've seen that methods within a disciplinary matrix are epistemically reliable just to the extent that background conceptions are accurate enough and the inferential architecture reliable enough that projectibility judgments can reliably guide theory invention and theory testing. The required accuracy of background conceptions and the required reliability of inferential practices in turn depend on the deployment of terms which refer to causally important natural kinds. The basic lesson here is that the reliability of scientific practices (when and to the extent that they are reliable) depends on many dimensions of accommodation between theories, concepts, classificatory practices, inferential standards, features of experimental design, and so on; and also on the causal powers of the phenomena under study. The claim, about any scientific discipline, that its methods are epistemically reliable with respect to a given range of questions is always an empirical hypothesis not only about the subject matter of the discipline but about a variety of complex cognitive, social, linguistic, and classificatory practices.

3. IMPLICATIONS: EPISTEMOLOGY

3.1 Cognitive Architecture, Social Structures, and the 'Context of Invention'

When scientific practices are epistemically reliable, their reliability rests on its being the case that, often enough, among the projectible answers to a scientific question that are actually proposed and publicized there is one that is relevantly close to the truth. It is not sufficient that a relevantly approximately true theory would ideally be recognized as projectible given the current standards. Nor is it sufficient that somewhere or other there is a researcher who recognizes it as projectible. The lone, unfunded, unpublished researcher whose conjecture gets it right makes thereby no contribution to the reliability of scientific practices. What's required to establish epistemic reliability in a particular research domain is that the existing social, economic, political, and cultural factors be such that, often enough, an approximately true answer to a question within that domain will be publicized and that, often enough, research investigating it will be funded. One therefore cannot fruitfully distinguish between a posteriori questions about the 'context of invention' and allegedly a priori questions about the 'context of confirmation'. Only where suitably accurate theories are actually invented are the methods deployed in the context of confirmation epistemically reliable.

3.2 Epistemology and the Political Economy of Science

There are, of course, domains in which actual scientific practices do, often enough, generate approximately correct answers. Then there are domains where they do not. Sometimes, of course, that's because the research questions are too hard, given the available conceptual and theoretical resources, but in other cases the imagination of researchers—and their ability to publish or be funded—is constrained by political factors: by the role of political ideology in science. It's tempting to think that only in the latter sorts of cases are distinctly political factors at work in determining the extent of the reliability of scientific practices.

This is a mistake. Of course when powerful interests are served by the widespread acceptance of false theories that rationalize the status quo, social and economic constraints on theory invention (and promulgation) operate against the promulgation of more accurate alternatives. But in those domains where theory invention and promulgation do tend to support reliable practices, the explanation is, almost always, that these are domains where powerful interests are served by the acceptance of accurate theories. The exceptions are cases in which the (usually temporary) effects of oppositional movements lead to the generation and promulgation of critiques of prevailing ideology. In all cases, the question of whether practices within a domain are reliable is an a posteriori question, not just about the subject matter in that domain but about its political economy. (For further discussions, see Boyd 1999b, 2001a.)

3.3 Anti-foundationalism

Ordinarily one thinks of foundationalism with respect to a subject matter as the claim that there is some class of epistemically privileged statements such that whenever some proposition about that domain is known, then it would always be possible (perhaps just 'in principle') to justify that proposition from true premises in the privileged class. Any formulation along these lines tacitly presupposes that the inferential principles underwriting the justification in question are themselves justifiable either a priori or on the basis of premises in the privileged class. Epistemic privilege here is something like immunity from some sort of doubt. Foundationalism comes in more or less modest forms. One might require of foundational statements that they are utterly immune from rational doubt. More modestly, one might count as foundational statements that could be rationally doubted only as an exercise in an epistemology class. Modest foundationalism of the latter sort about scientific and everyday empirical knowledge is extremely plausible. It amounts to the idea that scientific knowledge is something like common sense iterated.

It is false. There are no inferential principles sufficient for empirical inquiry which can be justified on the basis of premises which could be doubted only as a philosophical exercise. Any pattern of inference from empirical premises (foundational or not) to general conclusions is justified only on the assumption that the relevant cognitive, classificatory, linguistic, and conceptual practices as actually implemented are suitably accommodated to the causal structures of the relevant subject matter. That's never something that could be doubted only as a philosophical exercise. We may be justifiably more or less confident about some patterns of inferences in some disciplines with respect to some questions, but the complexity both of the subject matters of the sciences and of the political economy of scientific practices rules out any version of foundationalism worthy of the name.

There is no special foundational stance from which philosophers can approach the epistemology of empirical knowledge. At least as regards empirical knowledge, epistemology is just one among the empirical sciences, just as philosophical naturalists have maintained. (For further discussions, see Boyd 1989, 1992, 1999b, 2001a.)

4. IMPLICATIONS: THE METAPHYSICS OF NATURAL KINDS

4.1 Natural Kinds are Social Constructions

When appeals to 'causal' or 'naturalistic' theories of reference and of natural kinds first gained prominence as tools available to scientific realists, the two key examples were the well-confirmed claim that water = H_2O and the speculative claim that pain = C-fibre firing. Each of these claims represents the sort of revisionary (in the sense of non-analytic) claim whose potential confirmability the 'causal' and 'naturalistic' conceptions were supposed to underwrite. In neither of these cases was there any supposition that the identified entities were fundamental features of the universe like elementary particles, fundamental fields, or whatever. In particular, the materialist scientific realists who took seriously the a posteriori excursions into metaphysics that their theories of reference and of kinds legitimized insisted that the naturalness of a natural kind was discipline relative: pains probably constitute a natural kind in psychology or neurophysiology but not in physics. They defended a non-reductive but materialist metaphysics.

When Putnam abandoned 'metaphysical realism' in favour of a broadly pragmatist approach (see, e.g., Putnam 1981, 1983) he assigned to metaphysical realists the conception that there is a single interest and discipline independent family of fundamental natural kinds and a single true (reductive) theory about them. He held that non-reductionist materialism, understood metaphysically, was an incoherent position. Strangely, this one-true-theory, one-true-ontology conception of realism—which Putnam introduced in order to refute scientific realism, understood as a metaphysical position—has figured prominently in the thinking of some more recent scientific realists (see, e.g., Ellis 2001, 2002; Psillos 1999).

What consistently developed scientific realism does imply is the very opposite metaphysical conclusion (Boyd 1989, 1999b). Locke maintained that while Nature makes things similar and different, kinds are 'the workmanship of men'. Gender bias aside, he was right to say this. Of course his conception of how that workmanship operated was, for the most part, empiricist and conventionalist rather than accommodationist (but see Shapiro 1999). Still, the lesson we should draw from the accommodationist conception is that the theory of natural kinds *just is* (nothing but) the theory of how accommodation is (sometimes) achieved between our linguistic, classificatory, and inferential practices and the causal structure of the world.3 A natural kind is nothing (much) over and above a natural kind term together with its use in the satisfaction of accommodation demands. ('What else?' you ask. Well, there's whatever is necessary to accommodate translations which preserve satisfaction of accommodation demands and to accommodate phenomena like reference failure and partial denotation.) Or, better yet, the *establishment* of a natural kind *consists solely in* the deployment of a natural kind term in satisfying the accommodation demands of a disciplinary matrix. Given that *the* task of the philosophical theory of natural kinds is to explain how classificatory practices contribute to reliable inferences, that's all the establishment of a natural kind could consist in. Biological taxonomists sometimes speak of the 'erection' of higher taxa, treating such taxa as, in a sense, human constructions. They are right—and the same thing is true of natural kinds in general. Natural kinds are social constructions: they are the workmanship of women and men.

4.2 Reality and Mind Independence

According to the accommodationist conception, natural kinds are social constructions suited to the inductive/explanatory projects of particular disciplines. In a certain sense they are mind, interest, and project dependent. Does this compromise their reality so that realist accommodationism actually amounts to an anti-realist project? This worry actually comes in two flavours.

Natural Kinds as Secondary Qualities (Or Something Like That).

Since at least the seventeenth century philosophers have sometimes been tempted to think that properties defined in terms of human responses, interests, or projects have diminished ontological status. We don't seem to have similar ontological qualms about properties defined in terms of other sorts of animals. Consider again the semantics of alarm calls in Belding's ground squirrels. It's perfectly good science—and perfectly good metaphysics too—to say that the two different sorts of calls refer to aerial predators in the one case and to terrestrial predators in the other. But, of course, by 'aerial predators' one here means aerial predators on ground squirrels (not, e.g., on insects), and similarly by 'terrestrial predators' one means terrestrial predators on ground squirrels. Each category of predators is defined in terms of causal relations to Belding's ground squirrels. No one would maintain that this makes them irreducibly *rodential* and thus diminishes their ontological standing. Ground squirrels are real, as are aerial and terrestrial predation on them, so there's nothing metaphysically suspect about the two categories of predator in question. We're equally real, so kinds defined in terms of our causal capacities and responses are OK too. Fair play for humans!

Social Constructivism?

Here's a second worry. Might the accommodationist's conception that natural kinds are social constructions reflect an abandonment of scientific realism in favour of some sort of neo-Kantian social constructivism of the sort apparently defended by Kuhn (1970)?

No. What's distinctive about the view that Kuhn seemed to advocate is the idea that somehow or other the adoption of a paradigm imposes (successfully) a causal structure (one consonant with the fundamental laws of the paradigm) on the phenomena under study. Nothing in the accommodationist conception entails that the adoption of a research paradigm has any impact on the causal structure of the world beyond that caused, in ordinary ways, by the human practices involved. Indeed, the whole point of the accommodationist conception is that human conceptual and inferential practices must be accommodated to the causal structures of the phenomena under study, not *vice versa*. Accommodationism thus endorses the key realist, anti-constructivist claim that human social practices make no non-causal contribution to causal properties and relations in the world (see Boyd 1990).

4.3 'Reality' One More Time

What then are we to make of questions about the 'reality' of particular natural kinds? Surely such questions often raise some sort or other of legitimate ontological concerns. For example, some critics of racism in science deny the reality of human races as they are currently understood. How should their position be understood?

I suggest that questions about the reality of (alleged) natural kinds should always be understood as questions about the suitability of those kinds for induction and explanation *in particular disciplinary matrices*. The critic who denies the 'reality' of races would then be understood to be denying that races, as currently understood, play an epistemically legitimate role in biology. She would not then need to deny that those very categories are natural kinds in the social sciences that study stratification, poverty, and political oppression. (For a useful discussion of the relevant science, including issues about whether or not some racial categories might be, as one might say, 'natural kinds in medicine and pharmacology', see Social Science Research Council 2007.)

Can this approach be faulted on the grounds that it makes questions about the reality of kinds somehow mundane rather than metaphysical? No. If questions about the systematic relations of our epistemic practices to causal structures are not metaphysical questions, then it's not a metaphysical fact that water is H_2O. Of course, the realist naturalist's conception of 'reality' questions is less elevated than other conceptions might be, but that's the fate of naturalistic metaphysics.

5. IMPLICATIONS: REFERENCE

5.1 A Picture and Three Problems

According to naturalistic conceptions, the reference relation between a natural kind term, *t*, and a kind *k* is a matter of epistemically relevant causal relations connecting uses of *t* with instances of *k*. This suggests a metaphysical picture: reference is a relation between linguistic entities and entirely *extra-linguistic* (and in that sense *independently existing*) natural kinds. Natural kinds are, somehow or other, in the world, and available for discovery and naming, independently of human practices.

If one accepts this picture of semantic naturalism as many philosophers (including, e.g., Putnam 1978, 1980, 1983) seem to do, then it is subject to three challenges.

First, if we think of natural kinds as things somehow independent of linguistic and methodological practices, then there are lots of natural kinds out there, and it is difficult to see how the causal conception of reference fixing could explain how a natural kind term could ever have a unique referent (this seems to be the basis of the 'model theoretic' arguments in Putnam 1978, 1980).

This problem is exacerbated by the fact that causal theories of reference arose as criticisms of *descriptivist* (and conventionalist) conceptions of reference in the empiricist tradition. This has led some defenders of causal theories, and some critics (e.g., Putnam 1978, 1980), to conclude that naturalistic conceptions of reference, *at least if they are to underwrite correspondence conceptions of truth*, must be *pure causal theories* in the sense that they do not invoke descriptions or other conceptual elements, like referential intentions, in explaining reference. On this conception challenges regarding determinateness will seem even more acute.

Finally, reference to natural kinds is supposed to explain the inductive successes of scientific practice, so there must be some quite intimate connection between natural kinds and the conceptual machinery of the sciences. If one thinks of naturalistic theories as entailing that natural kinds are independent of that machinery, it is hard to see how the explanation could work unless it rested on some sort of *objective idealist* theory according to which natural kinds are somehow metaphysically 'fitted' for explanation and induction independently of the relevant practices. But, that's not consistent with any sort of naturalism. So, as Putnam and defenders of the 'natural ontological attitude' like Fine (1984) suggest, naturalistic correspondence theories seem to be in trouble.

The accommodationist conception avoids all these difficulties. Kind definitions and reference are aspects of the very same phenomenon of accommodation of conceptual and linguistic practices within a discipline to relevant causal structures, so there is no issue of how terms come to refer to free-standing natural kinds. Nor is there a problem explaining how independently existing kinds come somehow to be fitted to underwrite induction and explanation. They're not 'independent' of inductive and explanatory practices but aspects of them. Finally, the accommodationist conception is not a 'pure' causal conception in the sense indicated. All of the causal relations that figure in the satisfaction of the epistemic access and accommodation conditions are concept, interest, and language involving, and the very subject matter of the theory of natural kinds is the explanation of various human cognitive and practical achievements.

5.2 Descriptive, Conceptual, and Intentional Factors are Treated as Causal Factors

The accommodationist conception *entails* that descriptive, conceptual, and intentional factors figure fundamentally in establishing reference to natural kinds—and to establishing kind definitions, since these are the same phenomena according to accommodationism. Both the epistemic access and accommodation conditions make reference to such phenomena. According to accommodationism, a tendency toward (approximately) truthful actual *predication* of natural kind terms in service of (at least some of) the *aims and intentions* of participants in disciplinary matrices is fundamental to reference. So reference (and kind definition) are essentially concept, description, and intention involving.

Nevertheless, Reference is Not a 'Causal-Descriptive' Phenomenon.

Despite the deep involvement of descriptive and intentional factors in the accommodationist conception of reference, it would be highly misleading to describe it as a hybrid 'causal descriptive' conception. The conceptual, descriptive, and intentional factors which, according to accommodationism, figure in reference and in kind definition are all to be understood as causal factors in the relevant cognitive and social practices and in their engagement with the world. The accommodation condition, for example, refers to actual aims and intentions which play a causal role in disciplinary practices; the descriptive aspects of reference are a matter of the ways in which the actual deployment of descriptive resources in the relevant community and in the cognitive architecture of their members contribute (*causally*) to the epistemic reliability of practices; epistemic access is a matter of the extent to which the actual properties of things in the world causally regulate actual predications.

So, the accommodationist conception differs from other causal conceptions of reference not by adding *extra-causal* conceptual, descriptive, or intentional factors to the characterization of reference (and kind definitions) but by emphasizing the causal role of actual conceptual, descriptive, and intentional practices.

5.3 Two-dimensional Conceptions of Reference

According to two-dimensional conceptions (see, e.g., Chalmers 2004, 2005), one can ask, about a term *t*, used in one possible world *w*, two different questions regarding a different possible world *w'*. One can ask, about *t* used in *w*, what entity, if any, in *w'* it refers to. One can also ask about *t* what it would refer to in *w'* if were used in *w'* as it is used in *w*. The idea is that one can export (so to speak) from *w* to *w'* the reference-fixing machinery for *t* and then ask what that machinery would pick out in *w'*.

The conceptions of reference that are most congenial to two-dimensional approaches are ones which resemble the conception criticized in section 5.1. Suppose that one (a) thinks of natural kinds as somehow 'given' and individuated across possible worlds independently of discipline-specific linguistic and epistemic practices, and (b) thinks of the use in practice in *w*, of each natural kind term, *t,* as defined by its own independent causal profile which makes its use in *w* somehow latch onto the natural kind to which it refers in *w*. Then it would be easy to see how to understand the question 'what would *t* refer to in *w'* if it were used in *w'* as it is used in *w*?' One would imagine exporting the distinctive use profile of *t* in *w* to *w'* and asking which, if any, of the already given natural kinds that profile latched onto in *w'*. The only serious methodological question would be what the appropriate individuation conditions are for use profiles. One would probably not want, for example, to include the causal connection to *t*'s referent in *w* as constitutive of its use profile.

On the accommodationist conception of natural kinds and reference answering the question 'what would t refer to in w' if it were used in w' as it is used in w?' is much more difficult. Three features of the accommodationist conception ensure that this is so.

(1) Reference *and kind definitions* are determined simultaneously for all the terms deployed in a disciplinary matrix by facts about the linguistic and methodological practices in that disciplinary matrix. So what one would have to 'export' from w to w' would have to be the linguistic and methodological practices of the disciplinary matrix in which t is deployed. Complex issues would thus be raised about the individuation of disciplinary matrices.

(2) According to accommodationism the theory of reference and of natural kinds figures in the explanation of inductive and explanatory successes. Once the disciplinary matrix in which t is deployed has been 'exported' to w', determining the referent of t in w' would not be a matter of asking to which, if any, of the practice-independent natural kinds the use of t in w' would afford epistemic access. Instead, for there to be any determinate answer (other than 'does not refer'), it would have to be the case that the exported disciplinary matrix exhibits in w' inductive and explanatory successes very much like those achieved by that matrix in w. For worlds w' substantially different from w there would be no reason to expect that this would be so.

(3) According to accommodationism, the real definitions of natural kind terms, and thus the reference relation, are determined by the (epistemic reliability enhancing) features of actual linguistic, classificatory, and inferential practices. The theory of reference and of natural kinds is the theory of how (to what extent, and in what respects) the conceptual and inferential practices within a discipline recruit the cooperation of the world in achieving inductive and explanatory successes.

The nature and the stability of such practices thus depends not only on the conceptual resources, theoretical commitments, and inferential inclinations of the relevant practitioners but also on the actual reference-grounding causal interactions they have with their subject matter. For example, inferential practices within a disciplinary matrix tacitly presuppose that, often enough, an approximately satisfactory answer to a scientific question will be among those diagnosed in the matrix as projectible. Likewise they presuppose that, often enough, different independent strategies for detecting or measuring or otherwise assessing the phenomena under study will yield converging results; otherwise the methodological preference for seeking independent tests for proposed hypotheses would be unmotivated. The fact that practices within a matrix guided by these presuppositions tend to yield coherent conclusions—the fact that they enhance rather than destabilize the stability of the relevant methodological practices—is explained, according to accommodationism, by the reference constituting causal relations connecting the use of kind terms within the discipline to their referents; that is, by the satisfaction of the epistemic access and accommodation conditions.

What does this imply about how we should understand the exportation of a disciplinary matrix from *w* to *w'*? First, the 'exportation' won't have been successful—the exported practices won't have the stability and epistemic successes characteristic of a disciplinary matrix—unless the epistemic access and accommodation conditions are satisfied in *w'* by an assignment of terms to referents in *w'*. On any plausible way of individuating disciplines, the exported disciplinary matrix won't be the same matrix as that in *w* (rather than some other disciplinary matrix with terms exhibiting a similar orthography or phonology) unless the stable methodological practices in the exported matrix are very similar to those of the matrix in *w*. As we have seen, that stability has to be underwritten by the causal powers of the referents, in *w'*, of the relevant natural kind terms. So the natural kinds to which the relevant terms refer in *w'* must be very similar in their causal profiles to those to which the corresponding terms refer in *w*.

Conclusion: The successful exportation of a disciplinary matrix is possible only when *w'* is very nearby *w* (or perhaps, on an alternative conception of the individuation of properties and causal powers, only if *w'* is either nearby *w* or isomorphic to a world nearby *w*). The question 'what would *t* refer to in *w'* if it were used in *w'* as it is used in *w*?' has a determinate answer (other than 'does not refer') only for cases in which *w* and *w'* are very similar. The ease with which we intuitively address the question 'to whom would the name "Gödel" refer in a world in which Kurt Gödel didn't prove the incompleteness theorem?' masks real difficulties in nontrivial applications of two-dimensionalism with respect to natural kind terms.

5.4 Conceptual Role Semantics

Some especially important versions of two-dimensionalism maintain something like the following: (1) Associated with every natural kind term is an inferential or explanatory role accessible by conceptual analysis for the scientist who is fully conceptually competent. (2) The referent of that term will be whatever natural kind most nearly satisfies the conceptually given inferential/explanatory role. Accounts of this sort are easily generalized (perhaps must be generalized) to accommodate the point made earlier that reference is determined simultaneously for all the terms in a disciplinary matrix. The idea would then be that conceptual analysis will identify the different interrelated inferential and explanatory roles associated with the (most important?) terms within a disciplinary matrix and the referents of those terms will be given by whatever assignment best satisfies the inferential and explanatory roles picked out by conceptual analysis.

Of course versions of this sort of two-dimensionalism entail that we can achieve some substantial knowledge of the definitions of natural kinds by a sort of conceptual analysis that's pretty close to being *a priori*, so they're at odds with the broadly naturalistic conception that philosophy is a largely (or completely) *a posteriori* discipline continuous with the empirical sciences. In order to appreciate the differences between the two approaches, we first need to examine their similarities. Conceptual role semantics has two important affinities with accommodationism—one completely obvious, the other only slightly less so. We'll need to examine each.

Reference and Inferential Role.

Each approach focuses on the suitability of natural kinds for inference and explanation. In this regard, they differ in a crucial respect. Conceptual role semantics portrays reference as being determined by what fully competent scientists *believe (perhaps only tacitly) to be* the inferential and explanatory roles appropriate to the kinds to which they refer. The inferential/explanatory roles uncovered by the analysis of the concepts of fully competent researchers reflect what they (perhaps only tacitly) *believe* they are accomplishing with respect to induction and explanation.

Accommodationism focuses instead on what scientists are *actually* accomplishing. Like the conceptual role approach, accommodationism implies that when a natural kind term refers, those who deploy it must be getting something systematically right, both in what they believe about their subject matter and in the inferential strategies they employ; that's what the epistemic access and accommodation conditions require.

Where the approaches differ is that conceptual role semantics implies that, among the many beliefs and inferential practices characteristic of a discipline, those which are most fundamental conceptually (as revealed by conceptual analysis) will always (assuming that the discipline has a subject matter) be sufficiently accurate (if beliefs) or reliable (if inferential practices) that they pick out the referents of its natural kind terms in the required way. By contrast, the accommodationist conception leaves open the possibility of cases in which the epistemic access and accommodation conditions are satisfied for the terms in a discipline—and reference thus established—even though conceptually most central beliefs and inferential practices in a discipline (those that would be revealed analyses of the concepts of fully competent practitioners) do not single our the referents of those terms.

When we ask, for example, what stars and planets ancient astronomical writings refer to, we do not ask—and we need not ask—what astronomical entities, if any, best satisfied the conceptual/explanatory roles that were most central to the concepts the authors associated with the referring expressions. Instead, we look for entities to which they had epistemic access and regarding which they had explanatory or (especially in this case) inductive successes *whether or not those successes corresponded to their most conceptually central inductive or explanatory expectations*. We do exactly what the accommodationist conception dictates. The same is true with respect to the reference of disease terms in ancient medical and alchemical texts. We don't doubt that alchemists succeeded in referring to some of (what we call) the elements like sulphur and mercury, even if there are no substances which satisfied most of the most central inductive and explanatory patterns they associated with the relevant terms.

Of course the accommodationist conception is compatible with the (pretty obviously correct) view that pretty often the conceptually central beliefs and inferential strategies in a discipline contribute fundamentally to its epistemic successes and (thus) to determining the reference of its kind terms. But it licenses our dissent from the central conceptual machinery even of recent and current scientific theories. We may hold that in the racist late nineteenth- and twentieth-century literature on intelligence differences between human populations, terms for races referred to distinct biological populations even if we also hold that the most conceptually central inferential and explanatory practices in that literature were so flawed that there are no human populations which 'best fit' those practices because no human populations fit them at all.

Reference Communication and Conceptual Meanings.

There remains, however, another way in which conceptually central beliefs and inference patterns play a role in reference that we need to consider if we're to fully appreciate the similarities and differences between conceptual role semantics and accommodationism. Much of the plausibility of conceptual role semantics rests on the idea that in any given discipline there are beliefs and inference patterns, many of them tacit, such that someone ignorant of them is not fully competent to understand the literature in the field or to understand and appreciate the methods deployed in it. Someone ignorant of these commitments could, of course, use the terms in question with their standard referents, but she would have to do this by relying on the competence of experts (by 'borrowing' reference from them). It is these commitments that would be uncovered by an adequate analysis of the concepts of fully competent experts.

Although, strictly speaking, someone could subscribe to the accommodationist position offered here while failing to recognize the truth of this claim, a fully articulated accommodationist conception would surely entail it. Remember that, according to accommodationism, reference is a matter of socially coordinated epistemic access to kinds, reference to which plays a role in the epistemic successes (such as they are) in inductive/explanatory projects. So, whatever commitments, substantive or methodological, underwrite the possibility of reliable communication within a scientific discipline play an important *causal* role in underwriting the reference relations between the kind terms in that discipline and their referents. In fact, there are, in all technical disciplines, substantive and inferential commitments, many of them tacit, which are presupposed in such a way that ignorance of them would prevent someone from understanding the relevant theories and issues and from understanding the literature. (This is a main point of Kuhn 1970, and is completely independent of Kuhn's position about reference. See Boyd 1992, 2001a.) Let's call commitments which figure this way in the intelligibility of the professional discourse involving a term it *conceptual* meaning. Then, fully articulated accommodationism will agree with inferential role semantics that the conceptual meanings which would be uncovered by the analysis of expert concepts are centrally important in explaining reference.

The issue between these conceptions is thus not about whether or not conceptual meanings, so understood, figure in the establishment of reference. Instead, it is about how conceptual meanings contribute to reference. The accommodationist conception leaves open the possibility that the commitments reflected in the conceptual meaning of a term can be very seriously misleading, so that the epistemic access and accommodation conditions for the term are satisfied *despite* rather than *because of* those commitments. Meaning such as this would be *malignant*.

Some terminology will help to clarify the notion of malignant meanings. The conceptual meanings of terms in a discipline at a time are constituted by those substantive and inferential commitments such that ignorance of them prevents one from understanding the literature and inferential practices in the discipline at that time. Of course one need not subscribe to the relevant commitments in order to avoid ignorance. Certainly there are conceptual meanings in all systematic scholarly disciplines. Consider the situation of a historian of theology examining the theological literature of a particular religious tradition. In order to understand that literature she must, let us say, knowledgeably *engage with* the meaning-constitutive substantive commitments and inferential practices of that literature. She need not, of course, accept those commitments and practices. Atheists can do history of theology; monotheists can do the history of theology of polytheistic traditions, and *vice versa*.

Let's recognize two sorts of engagement:

Uncritical engagement: The sort of acceptance of cognitive/inferential commitments which characterizes ordinary fully competent participation in a research tradition at a time.

Critical engagement: The special stance which a historian or philosopher, or a methodologically self-conscious participant, might adopt (perhaps at some later time, or in response to considerations not currently reflected in the practices in the tradition, or in response to extra-disciplinary critiques) regarding the paradigmatic substantive and inferential commitments of the tradition without necessarily herself accepting the relevant commitments.

What conceptual role semantics presupposes is, roughly, that uncritical engagement gets the nature of the kind in question roughly right. What the accommodationist conception holds open is the possibility that sometimes uncritical engagement gets things deeply wrong in ways not diagnosable by fully competent practitioners. Often deeply erroneous commitments in a scientific discipline at a time can only be corrected in the light of discoveries—either within the discipline at a considerably later time or in some other discipline(s), often the social history of the discipline itself. In such cases the meanings revealed will be *malignant* rather than benign. This is surely so in some cases of some ancient sciences. Cases involving deeply entrenched racial or gender stereotypes in the history of the biology and psychology of human cognitive abilities and behavioural dispositions almost certainly provide real-life examples. So does current research in human sociobiology, as we'll see in the following 'worked example'.

Malignant Meanings in Extrapolative Human Sociobiology

Recent work in the discipline of human sociobiology (or 'evolutionary psychology') provides a good illustration of the role of malignant meanings in contemporary science. I'll here summarize briefly the case that this is so (for details of the argument for malignancy see Boyd 2001a; for other critiques of the extrapolative trend in human sociobiology see, e.g., Kitcher, 1985, Buller 2005).

By *extrapolative* human sociobiology I have in mind the research strategy which is grounded in the idea that findings from evolutionary theory provide independent constraints on theories of human developmental psychology, so that some theoretical issues can, at least *prima facie,* be resolved by appeals to 'predictions' from evolutionary theory (for a spirited defence see Alcock 2001). To a very good first approximation the central inferential patterns in human sociobiology involve (1) advocating an evolutionary scenario, *S*, regarding selection for a behaviour, *B*, in the environment of evolutionary adaptation and then (2) taking that scenario to 'predict' that humans have an innate and relatively nonmalleable unconscious motive with the same propositional content as the evolutionary function which S assigns to B. (For an even better approximation, add some inference patterns which trade on conflating the psychological use of 'altruism' and 'altruistic' with technical metaphorical uses of those terms in evolutionary theory. For an almost perfect approximation, add inferences from premises of the form '*B* has a biological/genetic basis' to '*B* is innate and relatively nonmalleable' [see Kitcher 1987; Buller 2005; Boyd 2001a].)

These inference patterns reflect deep confusions about the evolution of behavioural repertoires and about the relationship between evolved behaviours and learning. No evolutionary biologist would admit to accepting them if they were made explicit (see Alcock 2001 for repeated denials that contemporary human sociobiology has methodological commitments like these). Nevertheless, one cannot understand the literatures—one cannot see what inferential connections are being taken for granted—unless one engages with these pathologically defective inference patterns. They're malignant. Extrapolative human sociobiology studies, among other things, human mate choice, child rearing, child abuse, altruism, cooperation, and competition. Each of the terms 'human mate choice', 'child rearing', 'child abuse', 'altruism', 'cooperation', and 'competition', as they are used in the sociobiological literature, refer to real aspects of human psychology or behaviour, but they do so despite, rather than because of, the inferential strategies associated with their conceptual meanings in sociobiology.

The phenomenon of malignant meanings—in contemporary as well as ancient science—undermines completely the basic assumptions about concepts and reference which underwrite conceptual role semantics. Of course one might reinterpret conceptual role semantics so that it maintains that the referents of natural kind terms are determined by what conceptual analysis would reveal if it were informed by all the relevant facts about the discipline in question, its subject matter, and the reliability and unreliability of its actual practices. But this is just accommodationism.

6. IMPLICATIONS: METAPHILOSOPHY

6.1 *A Priori* Methods in Philosophy

With respect to the epistemology, semantics, and metaphysics of inquiry into matters of fact, there are two reasons to suppose that proper philosophical methods will have a substantial *a priori* component. First, it is plausible that philosophers' judgments on matters of basic epistemology reflect *a priori* knowledge of the basic principles of legitimate inductive inference. Second, it is plausible that broadly *a priori* conceptual analysis provides knowledge of the fundamental definitions of natural kinds. Neither of these plausible views is true. The epistemology, semantics, and metaphysics of scientific (and everyday) inquiry are matters of *a posteriori* empirical inquiry continuous with the (other) empirical sciences, just as Quine suggested.

6.2 Relevant Related Sciences

The idea that all or most of philosophy is continuous with the empirical sciences has often gone with the idea that the most closely related sciences will be physics (for its contribution to our understanding of metaphysics) and individual psychology (for its contribution to naturalistic epistemology). If the position developed here is right, this view is much too restrictive. There are ontologically respectable natural kinds in all the scientific disciplines, so physics has no special priority in metaphysics. The reliability of scientific methods within a discipline, when they are reliable, depends on social, political, and economic factors as well as on the causal properties of the phenomena that constitute its subject matter. In cases of malignant meanings associated with political ideology, philosophical critiques are sometimes made possible not mainly by developments within academic disciplines but by political struggles against the sorts of oppression which political ideology rationalizes (see Boyd 1999b, 2001a). Successful philosophy is thus continuous not only with many other sciences but (sometimes) with progressive political struggles.

6.3 Philosophers, Intuitions, and 'Conceptual Analysis'

If the positions developed here are correct, then philosophers' intuitions about matters epistemological, semantic, or metaphysical and the results of their conceptual analyses are not sources of *a priori* knowledge. Instead, they're reflections of trained judgments just as are the intuitions or conceptual analyses of scientists, or historians or whoever. That doesn't make them useless. Much good science depends on trained judgments. Moreover, it is likely that philosophers are often especially good at certain sorts of conceptual analysis. Still, it matters a lot that we recognize the *a posteriori* character of such judgments and, especially, that we recognize how wide a range of sciences and practices (including political ones) our work is continuous with.

6.4 Philosophical Naturalism

I've argued here for a broadly naturalistic approach to philosophical research connected with issues about knowledge and representation of matters of fact. For all I've said here, a non-naturalistic approach might be appropriate for others areas of philosophy. I don't believe that for a moment, but it is beyond the scope of this chapter to explore that issue. What I think is especially important is that the sort of philosophical naturalism which receives *prima facie* support from realist philosophy of science is profoundly non-reductionist. To approach philosophy as a science does not require that one's approach be narrowly scientific. Indeed, quite the opposite!

NOTES

1. Think of predicating t_i of something some expression, a, as predicating '... has a as a member' of t_i.
2. Actually several different but inadequately distinguished doctrines about species are all referred to as 'pluralism'. I have picked the one which best illustrates the sort of partial denotation I have in mind. Wilson 1999 contains excellent discussions (and bibliography) regarding species pluralism in its various forms.
3. Actually, I agree with the suggestion, implicit in Quine 1969, that the theory of natural kinds can be thought of as extending as well to the ways in which accommodation is achieved in non-human inductive and inferential systems.

REFERENCES

Alcock, J. (2001) *The Triumph of Sociobiology*. Oxford: Oxford University Press.

Boyd, R. (1983) 'On the current status of the issue of scientific realism', *Erkenntnis,* 19: 45–90.

———. (1985a) 'Lex orandi est lex credendi', in P. Churchland and C. Hooker (eds), *Images of Science: Scientific Realism Versus Constructive Empiricism*, Chicago: University of Chicago Press.

———. (1985b) 'Observations, explanatory power, and simplicity', in P. Achinstein and O. Hannaway (eds), *Observation, Experiment, and Hypothesis in Modern Physical Science*, Cambridge, MA: MIT Press.

———. (1989) 'What realism implies and what it does not', *Dialectica* 43, vol 1–2: 5–29.

———. (1990) 'Realism, conventionality, and "realism about" ', in G. Boolos (ed.), *Meaning and Method*, Cambridge: Cambridge University Press.

———. (1991) 'Realism, anti-foundationalism and the enthusiasm for natural kinds', *Philosophical Studies,* 61: 127–48.

———. (1992) 'Constructivism, realism, and philosophical method', in J. Earman (ed.), *Inference, and Other Philosophical Frustrations*, Berkeley and Los Angeles: University of California Press.

———. (1999a) 'Homeostasis, species, and higher taxa', in R. Wilson (ed.), *Species*: *New Interdisciplinary Essays*, Cambridge, MA: MIT Press.

———. (1999b) 'Kinds as the "workmanship of men": realism, constructivism, and natural kinds', in J. Nida-Rümelin (ed.), *Rationalität*, *Realismus*, *Revision*: *Proceedings of the Third International Congress*, *Gesellschaft für Analytische Philosophie*, Berlin: de Gruyter.

———. (2001a) 'Reference, (in)commensurability, and meanings: some (perhaps) unanticipated complexities', in P. Hoyningen-Huene and H. Sankey (eds), *Incommensurability and Related Matters*, Dordrecht, Netherlands: Kluwer.

———. (2001b) 'Truth through thick and thin', in R. Schantz (ed.), *What Is Truth?* Berlin and New York: Walter de Gruyter.

———. (2003a) 'Finite beings, finite goods: the semantics, metaphysics and ethics of naturalist consequentialism, part I', *Philosophy and Phenomenological Research*, 66: 505–53.

———. (2003b) 'Finite beings, finite goods: the semantics, metaphysics and ethics of naturalist consequentialism, part II', *Philosophy and Phenomenological Research*, 67: 24–47.

Buller, D. (2005) *Adapting Minds: Evolutionary Psychology and the Persistent Quest for Human Nature*. Cambridge, MA: MIT Press.

Chalmers, D. J. (2004) 'Epistemic two-dimensional semantics', *Philosophical Studies*, 118: 153–226.

———. (2005) 'The foundations of two-dimensional semantics', in M. Garcia-Carpintero and J. Macia (eds), *Two-Dimensional Semantics: Foundations and Applications*, Oxford: Oxford University Press.

Ellis, B. (2001) Scientific Essentialism. Cambridge: Cambridge University Press.

———. (2002) The Philosophy of Nature. Chesham, UK: Acumen.

Fine, A. (1984) 'The natural ontological attitude', in J. Leplin (ed.), *Scientific Realism*, Berkeley: University of California Press.

Goodman, N. (1973). *Fact, Fiction and Forecast*, 3rd edn. Indianapolis and New York: Bobbs-Merrill.

Hendry, R. F. (2005) 'Lavoisier and Mendeleev on the elements', *Foundations of Chemistry*, 7: 31–48.

Kitcher, P. (1985) *Vaulting Ambition: Sociobiology and the Quest for Human Nature*. Cambridge, MA: MIT Press.

Koslowski, B. (1996) *Theory and Evidence: The Development of Scientific Reasoning*. Cambridge, MA: MIT Press.
Kuhn, T. (1970) *The Structure of Scientific Revolutions*, 2nd edn. Chicago: University of Chicago Press.
Lipton, P. (1991) *Inference to the Best Explanation*. London: Routledge and Kegan Paul.
———. (1993) 'Is the best good enough?', *Proceedings of the Aristotelian Society,* 93: 89–104.
Psillos, S. (1999) *Scientific Realism*: *How Science Tracks Truth*. New York and London: Routledge.
Putnam, H. (1978) *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.
———. (1980) 'Models and reality', The Journal of Symbolic Logic, 45: 464–82.
———. (1981) *Reason*, *Truth and History*. Cambridge: Cambridge University Press.
———. (1983) 'Vagueness and alternative logic', in his *Realism and Reason*. Cambridge: Cambridge University Press.
Quine, W. V. O. (1969) 'Natural kinds', in his *Ontological Relativity and Other Essays*. New York: Columbia University Press.
Shapiro, L. (1999) 'Toward "perfect collections of properties": Locke on the constitution of substantial sorts', Canadian Journal of Philosophy, 29: 551–93.
Sherman, P. W. (1977) 'Nepotism and evolution of alarm calls', *Science*, 197: 1246–53.
Sherman, P. W. and Reeve, H. K. (1997) 'Forward and backward: alternative approaches to studying human social evolution', in L. Betzig (ed.), *Human Nature: A Critical Reader*, New York: Oxford University Press.
Social Science Research Council (2007) 'Is Race "Real"?' Online. Available HTTP: <*http://raceandgenomics.ssrc.org*> (accessed 15 September 2009).
Wilson, R. (1999) *Species*: *New Interdisciplinary Essays*. Cambridge, MA: MIT Press.

Contributors

Helen Beebee is Professor of Philosophy at the University of Birmingham, UK. Her publications include *Hume on Causation* (Routledge 2006) and *Truthmakers: The Contemporary Debate* (co-authored/edited with Julian Dodd, OUP 2005), and articles in *Mind*, *Noûs*, *The Journal of Philosophy*, *Philosophical Review,* and other journals.

Corine Besson is a fixed-term tutorial fellow in philosophy at St Hugh's College, University of Oxford. She works mainly on the philosophy of logic and language, and the epistemology of logic and language. Her publications on natural kind terms include 'Externalism, internalism and logical truth', *The Review of Symbolic Logic*, 2 (2009).

Alexander Bird is professor of philosophy at the University of Bristol. He is the author of *Nature's Metaphysics: Laws and Properties* (Oxford 2007), and has published a number of papers defending natural kinds essentialism.

Richard Boyd is Susan Linn Sage professor of philosophy at Cornell University. He works primarily in the philosophy of science, focusing on issues like natural kinds, reference, objectivity, and realism that lie in the intersection of metaphysics and epistemology. He also works in metaethics and philosophy of biology.

Robin Hendry is reader in philosophy at Durham University. His publications include 'Elements, compounds and other chemical kinds' in *Philosophy of Science*, 73 (2006) and *The Metaphysics of Chemistry* (forthcoming with Oxford University Press).

Joseph LaPorte is an associate professor of philosophy at Hope College, in Holland, Michigan. Publications relating to his contribution here include *Natural Kinds and Conceptual Change* (2004, Cambridge University Press) and 'Is there a single objective, evolutionary tree of life?', *The Journal of Philosophy* 102 (2005).

Genoveva Martí is ICREA research professor in the Departament de Lògica, Universitat de Barcelona. Her publications include 'Rigidity and general terms' (*Proceedings of the Aristotelian Society*, 104 [2004]). She has co-authored (with José Martínez-Fernández) 'General terms and non-trivial rigid designation' (C. Martínez, J. L. Faguera, and J. M. Saguillo [eds], *Current Topics in Logic and Analytic Philosophy*, Universidade de Santiago de Compostella 2007).

José Martínez-Fernández is associate professor ('professor agregat') in the Departament de Lògica, Universitat de Barcelona. His work on philosophical logic has been published in the *Notre Dame Journal of Formal Logic*, *Philosophical Studies,* and *International Journal of Philosophical Studies.*

Harold Noonan is professor of philosophy at the University of Nottingham. His publications include *Personal Identity* (Routledge 1989, 2nd edn. 2003), *Frege* (Polity 2001), *Guidebook to Hume on Knowledge* (Routledge 1999), and *Hume* (OneWord Publications 2007).

Nigel Sabbarton-Leary is completing his PhD at the University of Birmingham, UK. His PhD, tentatively titled 'Naming without necessity', addresses issues in the semantics and metaphysics of natural kinds. His publications include N. Leary, 'How essentialists misunderstand Locke', *History of Philosophy Quarterly* 26 (2009).

Emma Tobin is an AHRC metaphysics of science postdoctoral fellow at the University of Bristol. Emma has co-authored the entry 'Natural kinds' for the *Stanford Encyclopedia of Philosophy* (http://plato.stanford.edu) with Alexander Bird.

Åsa Wikforss is professor of theoretical philosophy at Stockholm University, Sweden. Her publications include 'Naming natural kinds', in *Synthese* 145 (2005), and 'Semantic externalism and psychological externalism', in *Philosophy Compass* 3 (2008).

Jessica Wilson is associate professor of philosophy at the University of Toronto. She has published on a variety of topics pertaining to the modality and metaphysics of broadly scientific entities.

Index