

Realism, Natural Kinds and Philosophical Methods

0. 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 metaphysical 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.

1. Realist philosophy of science

1.0. Projectibility and evidence. It’s 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 (*sensu* 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 (*sensu* 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 (= rival theories also projectible by those standards, and

3. E was gathered so as to control for artifacts suggested by theories projectible by those standards.

1.1. 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 (= 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 rules 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.

Basic idea: Theory dependent methods work just in case the relevant background theories and methods are accurate and reliable (see Boyd 1983, 1985a, 1985b, 1989).

1.2. Natural kinds and reference.

1.2.0. The basic accommodationist picture. Paul Sherman and his associates (Sherman 1977, Sherman and Reeve 1997) have identified two different sorts of alarm calls 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 behaviors 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 \underline{a} and \underline{t} ? It's by no means the case that \underline{a} s are invariably associated with aerial predators or \underline{t} s with terrestrial ones; there are false positives and false negatives. Instead, what makes \underline{a} and \underline{t} refer to aerial and terrestrial predators respectively is instead that (1) there is a tendency for \underline{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 \underline{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 a, b; 1991; 1999a,b; 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's 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.

1.2.1. 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 Ks explain such inductive and explanatory successes as we have achieved. H₂O 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 H₂O 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 disciplines inferential practices. Here's a more precise way of saying all this.

Let M be a disciplinary matrix and let t_1, \dots, t_n be the natural kind terms deployed within the discourse central to the inductive/explanatory successes of M. Then the families

F_1, \dots, F_n of properties provide definitions of the kinds referred to by t_1, \dots, 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 , $i=1, \dots, n$. 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, \dots, 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 requires 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).

1.2.2. 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 *sensu* 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 behavior, evolutionary biology, etc.--different notions of species are required, but that this need is

¹ . Think of predicating t_i of something some expression, a , as predicating "...has a as a member" of t_i .

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.

1.2.3. Homeostatic property cluster 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 the definition of a species would misrepresent biological reality and thus undermine accommodation. [For further discussions see Boyd 1999a, 1999b, 2003a, 2003b.]

1.2.3. 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).

1.3. 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 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.

2. Implications: Epistemology.

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.

2.0. 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.

2.1. 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.]

2.2. 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 utterly immune from rational doubt. More modestly, one might count as foundational statements that could be rationally doubted only as an exercise in a 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 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 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.]

3. Implications: The metaphysics of natural kinds.

3.1. 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 = H₂O and the speculative claim that pain = C-fiber 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 favor 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, 2003; 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 world³. 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.

3.1. 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 flavors.

3.1.0. Natural kinds as secondary qualities (or something like that). Since at least the 17th 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

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.

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 predators in question. We're equally real, so kinds defined in terms of our causal capacities and responses are OK too. Fair play for humans!

3.1.2. 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 favor 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).

3.2. "Reality" once more time. What then are we to make of questions about the "reality" of particular natural kinds? Such questions often raise some sort 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 H₂O. 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.

4. Implications: Reference.

4.0. 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, 1983b) 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 NOAers 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.

4.1. 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, No hybridization! 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.

4.2. Two-dimensional conceptions of reference. According to two dimensional conceptions 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 conception of reference which is most congenial to two dimensional approaches is one according to which each referring expression goes with its own independent reference fixing machinery.

On the accommodationist conception reference is established simultaneously for

all of the referring terms in a disciplinary matrix. Moreover, the factors which underwrite reference for those terms will involve not only complex conceptual relations between the uses of those terms but complex causal relations between features of the relevant subject matter, the uses of the terms in practice, and the political economy of the relevant disciplines. For this reason it's doubtful whether the two dimensional approach is applicable for natural kind terms except in cases where W' is *very* nearby W .

4.3. 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 above 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.

4.3.0. 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 out 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 were most central to the concepts the authors associated with the referring expressions. Instead we look for entities to which they had epistemic access are 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 sulfur 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 19th and 20th 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.

4.3.1. 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 s/he would have to do this by relying on the competence of experts (by “borrowing” reference from them). It is these commitments which 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 its 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's 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 an historian of theology examining the theological literature of a particular religious tradition. In order to understand that literature she must, let us say, knowledgably *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 fully competent participation in a research tradition at a time.

Critical engagement: The special stance which an historian or philosopher might—perhaps at some later time or in response to considerations not reflected in the practices in the tradition at that time—take regarding the paradigmatic substantive and inferential commitment 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, *t*, 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 behavioral dispositions almost certainly provide real life examples. So does current research in human sociobiology as we'll see in the following “worked example.”

4.3.2. 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 defense 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 behavior, *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 behavioral repertoires and about the relationship between evolved behaviors 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 behavior, 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.

5. Implications: Metaphilosophy.

5.0. *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's plausible that philosophers' judgments on matters of basic epistemology reflect *a priori* knowledge of the basic principles of legitimate inductive inference. Second, it's 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.

5.1. 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.

5.2. Philosophers, intuitions, and “conceptual analysis.” If the positions developed here are 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’s 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 judgments and, especially, that we recognize how wide a range of sciences and practices (including political ones) our work is continuous with.

5.3 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’s beyond the scope of this paper 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!

Richard N. Boyd
Cornell University

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.

Boyd, R. 1985a. Lex Orandi est Lex Credendi. in Churchland and Hooker (eds.) Images of Science: Scientific Realism Versus Constructive Empiricism. Chicago: University of Chicago Press.

Boyd, R. 1985b. "Observations, Explanatory Power, and Simplicity." In P Achinstein and O. Hannaway (eds.) Observation, Experiment, and Hypothesis In Modern Physical Science. Cambridge: MIT Press.

Boyd, R. 1989. "What Realism Implies and What It Does Not" Dialectica.

- Boyd, R. 1990. "Realism, Conventionality, and 'Realism About'" in Boolos, ed. Meaning and Method. Cambridge: Cambridge University Press.
- Boyd, R. 1991. "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds." Philosophical Studies 61: 127-148.
- Boyd, R. 1992. "Constructivism, Realism, and Philosophical Method." in John Earman, ed. Inference, and Other Philosophical Frustrations.
- Boyd, R. 1993. "Metaphor and Theory Change" (second version) in A. Ortony (ed.) Metaphor and Thought, 2nd Edition. New York: Cambridge University Press.
- Boyd, R. 1995. "Postscript to 'How to be a Moral Realist'" in Paul K. Moser and J.D. Trout, eds. Contemporary Materialism: A Reader. New York: Routledge.
- Boyd, R. 1999a. "Homeostasis, Species, and Higher Taxa," in R. Wilson, ed. Species: New Interdisciplinary Essays. Cambridge: MIT Press.
- Boyd, R. 1999b. "Kinds as the "Workmanship of Men": Realism, Constructivism, and Natural Kinds." in Julian Nida-Rümelin, ed. Rationalität, Realismus, Revision: Proceedings of the Third International Congress, Gesellschaft für Analytische Philosophie. Berlin: de Gruyter.
- Boyd, R. 2001a. "Reference, (In)commensurability, and Meanings: Some (Perhaps) Unanticipated Complexities." in Paul Hoyningen-Huene and Howard Sankey, eds. Incommensurability and Related Matters. Dordrecht: Kluwer.
- Boyd, R. 2001b. "Truth through Thick and Thin," in Richard Schantz, ed. What is Truth?. Berlin and New York: Walter de Gruyter.
- Boyd, R. 2003a. "Finite Beings, Finite Goods: The Semantics, Metaphysics and Ethics of Naturalist Consequentialism, Part I." Philosophy and Phenomenological Research, LXVI, No. 3 May 2003, 505-553.
- Boyd, R. 2003b. "Finite Beings, Finite Goods: The Semantics, Metaphysics and Ethics of Naturalist Consequentialism, Part II." Philosophy and Phenomenological Research, LXVII, No. 1 July 2003, 24-47.
Ithaca: Cornell University Press.
- Buller, D. 2005. *Adapting Minds: Evolutionary Psychology and the Persistent Quest for Human Nature*. MIT Press, Cambridge, MA.**
- Ellis, B., (2001), *Scientific Essentialism*, Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.
- Ellis, B., (2002), *The Philosophy of Nature*. Chesham: 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 edition. Indianapolis and New York: Bobbs-Merrill.

Hendry, R. 2005. "Lavoisier and Mendeleev on the elements." Foundations of chemistry 7:31-48.

Kitcher, Philip. 1985. *Vaulting Ambition: Sociobiology and the Quest for Human Nature*. Cambridge, Mass.: MIT Press.

Koslowski, B. 1996. *Theory and evidence: The development of scientific reasoning*. Cambridge:- MIT Press.

Kuhn, T. 1970. The Structure of Scientific Revolutions, 2nd edition. Chicago: University of Chicago Press.

Lipton, P. 1991. Inference to the Best Explanation. London: Routledge and Kegan Paul.

Lipton, P. 1993. "Is the Best Good Enough?" Proceedings of the Aristotelian Society 93/2: 89-104.

Mayr, E. 1969. Principles of Systematic Zoology. New York: McGraw-Hill.

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.

Putnam, H. 1980. "Models and Reality." The Journal of Symbolic Logic, Vol. 45, No. 3 (Sep., 1980), pp. 464-482.

Putnam, H. 1981. Reason, Truth and History. Cambridge: Cambridge University Press.

Putnam, H. 1983. "'Vagueness and Alternative Logic.'" in H. Putnam, Realism and Reason. Cambridge: Cambridge University Press.

Quine, W.V.O. 1969. "'Natural Kinds.'" in W.V.O. Quine, 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 (1999): 551-593.

Social Science Research Council 2007. "Is Race 'Real'?"
<http://raceandgenomics.ssrc.org/>