Draft. Published in Nida-Rümelin, ed. Rationalität, Realismus, Revision / Rationality, Realism, Revision, 52-89. ISBN 978-3-11-016393-3

Kinds as the "Workmanship of Men": Realism, Constructivism, and Natural Kinds[1]

0. Introduction

0.0. Natural Kinds and Their Essences: Enthusiasm and Critique. My topic is the theory of natural kinds. Let's begin by examining the current situation in philosophy and related disciplines with regard to natural kinds and with regard to related issues about realism and essentialism.

Within mainstream analytic philosophy, the "naturalistic" work of Putnam (1972, 1975a, 1975b) and Kripke (1971, 1972) on reference and essences, bolstered by naturalist and realist conceptions in the philosophy of science, has had the effect of establishing an almost unbridled enthusiasm for natural kinds and their essences. We confidently discern the metaphysical essences of individuals, states, properties, events, tropes--whatever--assessing the properties they have in "all metaphysically possible worlds." From philosophy of mind to aesthetics we deploy modal operators, understood metaphysically, with as much confidence as we do quantifiers.

Similarly, in much of analytic philosophy of science we confidently anticipate that the scientific phenomena we discuss will possess real rather than nominal essences, and our conceptions of everything from the nature of reduction to the commensurability of alternative paradigms is likely to be grounded in this judgment. Even contemporary empiricist philosophers of science often arrange to be realists about some sorts of natural kinds and their a posteriori real definitions. I'll argue later that, although naturalistic developments within realist philosophy of science contributed significantly to recent essentialist enthusiasms, the deployment of essentialist notions within scientific realism suggests a conception quite different from that which would be required to underwrite much of the enthusiasm elsewhere in the analytic tradition. For the present, however, it suffices to note that within mainstream analytic philosophy there is a flourishing of realism about natural kinds and their real essences.

By contrast, almost everywhere else in philosophy, and in the humanities generally, there have been very serious critiques of realism, and especially of essentialism. Even within the philosophy of science we have philosophers of biology and philosophically sophisticated biologists arguing that species (in contrast to, say, chemical compounds) aren't kinds at all, but individuals (Ghiselin 1974; Hull 1978), criticizing "essentialist thinking" in biology [The basic argument is due to Mayr; see, e.g., Mayr 1988, Hull 1965.], and (in the case of pheneticist and cladist conceptions) rejecting or trivializing the claim that higher taxa, like genera, families and orders, are natural kinds with real essences.

Much more severe are the criticisms offered by philosophers and others who affirm that

1. In formulating my approach to the issues discussed here, I have benefited greatly from conversations with Susan Babbitt, Alan Gilbert, Kristin Guyot, Ian Hacking, Eric Hiddleston, Karen Jones, Barbara Koslowski, Ruth Millikan, Satya Mohanty, Sydeny Shoemaker,

science--and other intellectual activities--are merely the "social construction of reality." At one end of the spectrum we have philosophers of science like Hanson and Kuhn who appear to defend a neo-Kantian and anti-realist conception. Somewhere nearby lie the anti-realist positions of Goodman and of the most recent temporal stages of Putnam, whose apparently relativist positions have much in common with Kuhn's. In the case of these anti-realist positions, the issue of natural kinds is crucially important. Putnam (1983) denies that there is a "ready made world" and his target is plainly what he takes to be a realist conception of natural kinds and categories. Like Goodman, he appears to think that we and "the world" jointly make the facts.

At the other end of the social constructivist spectrum come "postmodernist" thinkers in philosophy, literary theory, feminist theory, science studies and related disciplines. At least on ceremonial occasions, these thinkers appear to hold that to talk of "reality" is to participate in a discourse whose content is determined by complex relations of social and political power <u>rather than</u> by some putative phenomena in the world. For them, at least on ceremonial occasions, even talk about socially constructed <u>reality</u> is, strictly speaking, an unjustified concession to realism. Here too, there is often an especially sharp criticism directed at (what the thinkers in question call) <u>essentialism</u>, particularly regarding social and psychological kinds and categories.

0.1. <u>Stereotypes</u> <u>And</u> <u>Slogans</u>. Many of the critics of realism, and especially of realist conceptions of natural kinds and their essences, conceive of a realistic and naturalistic conception of natural kinds as entailing that natural kinds are:

1. independent of human practices.

2. defined by

    a. eternal,

    b. unchanging,

    c. ahistorical, and

    d. intrinsic

necessary and sufficient membership conditions;

3. referred to in

    a. fundamental,

    b. exceptionless,

    c. eternal, and

d. ahistorical

laws; and

4. discovered by the deployment of

    a. eternal,

    b. ahistorical, and

    c. foundational

scientific methods.

Although it is easy to find philosophers who defend a realist and essentialist conception of natural kinds but who dissent from one or more of the elements of this stereotype, the fact remains that it is a stereotype not a caricature: the elements mentioned are common to many realist essentialist conceptions of scientific knowledge generally, and of knowledge of natural kinds in particular.

In response to this picture, critics of realism and of essentialism have deployed certain characteristic slogans. They have variously maintained that natural kinds are (sometimes or always)

1. open textured,

2. historically situated,

3. relationally and historically defined, (and thus)

4. non-eternal, and
5. non-intrinsic

"social constructions" discovered (if that's the right term) by methods which are themselves

6. socially and historically situated,

7. irreducibly political, and

8. non-foundational

"social constructions."

My topic is this: What should we make of the conflict between realist essentialism and the position represented by these slogans? To what extent do we make kinds and other stuff?

What is the status of our knowledge of them?

0.2. <u>Credo</u>. My answer is, in brief, the following:

In the first place, a broadly realist and naturalist conception of natural kinds, causation, truth and knowledge is correct. Natural kinds do possess <u>a posteriori</u> real essences or natural definitions, as opposed to the nominal essences proposed by Locke and others in the empiricist tradition. Causation is not a social construction: we do not make causal relations, except in so far as we ourselves function as ordinary causal phenomena. Truth (about natural kinds, causal relations and the other fundamental subjects of science) is <u>correspondence</u> truth--socially constructed truth won't do. In so far as the knowledge of facts about the world is concerned, knowledge and rationality are matters of certain causally reliable tendencies towards approximately (correspondence-wise) true beliefs. Four cheers for realism!

On the other hand, that said, in <u>every</u> other respect <u>every</u> relativist and anti-realist sounding postmodernist slogan is (properly understood) true and important. Natural kinds are always, in an important sense, social constructions and practice relative. They are often not "eternal." Their defining properties are often neither intrinsic nor unchanging, nor need they determine necessary and sufficient defining conditions. They need not figure in exceptionless or eternal laws. The methods by which we learn about them are importantly historically situated, socially and politically constructed, and non-foundational. Moreover, each of these conclusions is a consequence of a properly developed naturalist, realist, and essentialist conception of kinds and of scientific knowledge.

0.3. <u>Strategy</u>: <u>Articulating A Really Realist</u>, <u>Really Naturalist</u>, <u>Theory of Natural Kinds</u>.

0.3.0. <u>Locke</u>. My main thesis is that a proper realist and naturalist response to constructivist and postmodernist conceptions of kinds requires the articulation of a much more thoroughgoingly realist and naturalist conception of kinds and kind terms. A surprising feature of that conception is that it incorporates--<u>as essential to naturalistic scientific realism</u>--a certain conception of the <u>practice dependence</u> of kinds. I believe that Locke was (gender bias aside) right to hold that, while Nature makes things similar and different, kinds are "the workmanship of men." He was also right, I hold, in attributing to Nature the making of causal similarity and difference. Kinds are practice dependent but (in a sense I'll make clear) <u>the world is not</u>. It will be a consequence of this account that properly developed naturalist realism about natural kinds and knowledge has the relativist and anti-realist sounding consequences just mentioned.

0.3.1. <u>Three Theses and Two Applications</u>. In articulating this "Lockean" version of realist naturalism, I'll identify, and briefly defend, three philosophical theses, two of whose further philosophical implications I'll then explore. They are these:

<u>Thesis One</u>: <u>The Accommodation Thesis</u>. This thesis is intended to capture the basic <u>realist</u> element in the naturalist realist conception of natural kinds: that their naturalness consists in a certain <u>accommodation</u> between the relevant conceptual and classificatory practices and <u>independently existing</u> causal structures, and that the achievement of knowledge of approximate <u>correspondence</u> truths is central to that accommodation.

Thesis Two: Anti-foundationalism. This thesis is intended to capture the claim that the ways in which accommodation depends on (approximate) correspondence truth entails a non-foundationalist conception of scientific methods--one according to which they do have the sort of historical, social and political situatedness beloved by postmodernists.

Thesis Three: The Bicameralism Thesis. This thesis is intended to make precise just what is true in the metaphysical claim that natural kinds are "social constructions." It serves to justify the claim mentioned earlier that the sort of essentialism properly inferred from naturalist realism in the philosophy of science does not fully underwrite the recent essentialist enthusiasms within analytic metaphysics.

I'll explore the implications of these three theses for our philosophical conceptions of essences, modality and "possible worlds," and for our understanding of the relation between politics and epistemology.

0.3.2. Keeping Score. Every so often, I'll pause to indicate the extent to which the arguments so far developed support the claim that the all of the stereotypical relativist sounding slogans about kinds and knowledge reflect important truths.

1. The Accommodation Thesis.

1.1.0. Accommodation and Reliable Induction.

The accommodation thesis is a thesis about natural kinds and about the philosophical theory of natural kinds. It is a truism that the philosophical theory of natural kinds is about how classificatory schemes come to contribute to the epistemic reliability of inductive and explanatory practices. Quine was right in "Natural Kinds" (1970) that the theory of natural kinds is about how schemes of classification contribute to the formulation and identification of projectible hypotheses (in the sense of Goodman 1973). It is likewise a truism that the naturalness of natural kinds consists in their aptness for induction and explanation, and that definitions of natural kinds are reflections of the properties of their members which contribute to that aptness.

The accommodation thesis makes the further claim that what is at issue in establishing the reliability of inductive and explanatory practices, and what representation of phenomena in terms of natural kinds makes possible, is the accommodation of inferential practices to relevant causal structures.

Here is the basic idea: Consider a simplified case in which reliable inductive practices depend on our having a suitable vocabulary of natural kind terms. Suppose that you have been conducting experiments in which you exposed various salts of sodium to flames. In each of many cases, the flame turned yellow. You conclude that always (or almost always) if a salt of sodium is heated in a flame, then a yellow flame results. You are right and your inference is scientifically respectable.

Your inductive success in this matter is, of course, a reflection of the fact that the categories salt of sodium, flame, and yellow are natural categories in chemistry, and of the fact that the hypothesis you formulated with the aid of reference to these categories is a projectible one.

Now anyone who has read Goodman can come up with indefinitely many unprojectible generalizations about such matters which equally well fit all past data but which are profoundly false. You were able to discern the true one because your inductive practices allowed you to identify a generalization which was appropriately related to the causal structures of the phenomena in question. In this particular case, what distinguished the generalization you accepted from the unprojectible generalizations which also fit the extant data was that, for any instantiation of it which makes its antecedent true, the state of affairs described by the antecedent will (in the relevant environment) <u>cause</u> the effect described by its consequent. Your deployment of projectible categories and generalizations allowed you to identify a <u>causally sustained</u> generalization.

What is true in this simplified example is true in general of our ability, in scientific practice, to identify true (or approximately true) generalizations: we can identify such generalizations just to the extent that we can identify generalizations which are (and will be) sustained by relevant causal structures. Things may be more hairy than they are in our example; perhaps the truth makers for the antecedents of true instantiations are effects of causes of the states of affairs described by the consequents. Perhaps the generalizations speak of causal powers and propensities rather than of determinate effects, so that it is the causal sustenance of propensities rather than the causation of effects which is relevant. Perhaps the generalizations have a more complex logical form, etc.

Still, we are able to identify true generalizations in science--and in everyday life--because we are able to accommodate our inductive practices to the causal factors which sustain those generalizations. In order to do this--in order to frame such projectible scientific generalizations at all--we require a vocabulary, with terms like "sodium salt" and "flame" which is itself accommodated to relevant causal structures. This is the essence of the accommodation thesis regarding theoretical natural kinds.

1.1.1. <u>Disciplinary Matrices and a Kind of Relativism</u>. It follows from the accommodation thesis that the naturalness of a natural kind depends on the inferential architecture within which representations of it are embedded. The kind <u>salt of sodium</u> is a natural kind just because reference to it contributes to the accommodation of the inductive and explanatory practices of chemists and others to relevant causal structures. Classifying reagents accurately as sodium salts and referring to them by the term "sodium salt" would make no such contribution except in the context of a whole bunch of theory dependent classificatory experimental and inferential practices involving--among other things--reference to lots of other chemical kinds.

In particular the accommodation thesis commends to us the terminology of philosophers who speak, for example, of psychological states like pain being natural kinds "from the point of view of psychology" but not (owing to multiple realizability, for example) "from the point of view of basic physics." Accommodation of the inferential practices of psychology to relevant causal structures requires descriptive resources like the term "pain," whereas accommodation of such practices in basic physics does not.

Thus the fundamental notion in the theory of theoretical natural kinds is not the notion of such a kind, simpliciter, but instead the notion of a kind's being natural with respect to a particular inferential architecture. When we talk simply of a natural kind, or of natural kinds generally, there is either tacit reference to some inferential architecture or tacit quantification over some domain of them. At least in the case of natural kinds in the sciences, that inferential architecture can best be thought of as being provided by a disciplinary matrix: a family of inductive and explanatory aims and practices, together with the conceptual resources and vocabulary within which they are implemented. The naturalness of a scientific natural kind is relative to the role reference to it plays in a disciplinary matrix.

1.1.1.0. Keeping Score, I: Mind and Practice Dependence. Here's an important sense in which natural kinds and their naturalness are not independent of human purposes, interests, aims and practices. If we adopt the standard realist and naturalist conception of natural kinds as vehicles for the identification of projectible generalizations, then practice dependence is entailed.

1.1.2. Accommodation Demands and "Accommodationism" Regarding Kind Definitions. Some terminology will prove useful. By the accommodation demands of a disciplinary matrix, M, let us understand the requirement of "fit" or accommodation between M's conceptual and classificatory resources and relevant causal structures which would be required in order for the characteristic inductive, explanatory (or practical) aims of M to be achieved. As I propose to use the term, there may be basically successful disciplinary matrices not all of whose accommodation demands can be satisfied: for some of the explanatory or inductive aims of such a disciplinary matrix there might not exist in the world the sorts of causal structures which could sustain the sought after generalizations or explanations.

The basic claim of the accommodation thesis is that the subject matter of the theory of natural kinds is how the use of use of natural kind terms and concepts (likewise for natural relation terms or natural magnitude terms, etc.) contributes to the satisfaction of the accommodation demands of disciplinary matrices, in so far as such accommodation is possible.

According to the position I am here developing there is a perfectly good sense of the term "definition" according to which a natural kind is defined by a certain causal role specified in terms of the inferential role which the use of a natural kind term referring to it plays in satisfying the accommodation demands of a disciplinary matrix. Call this sort of definition of a kind a programmatic definition. There is another perfectly legitimate sense of "definition" according to which a definition of a natural kind is provided by an account of the properties shared by its members in virtue of which reference to the kind plays the role required by its programmatic definition. Call this sort of definition of a kind an explanatory definition.

To a good first approximation [I'm ignoring here issues like partial denotation, non-referring expressions, etc.] I advocate the following "accommodationist" conception of kind definitions and of reference:

Let M be a disciplinary matrix and let $t_1,...t_n$ be the natural kind terms deployed within the discourse central to the inductive/explanatory successes of M. Then the families $F_1,..F_n$ of properties provide explanatory definitions of the kinds referred to by $t_1,...t_n$, 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,...n.

2. (Accommodation condition) This fact, together with the causal powers of things satisfying these explanatory definitions, causally explains how the use of $t_1,...t_n$ 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 and to obtain correct explanations: whatever tendency there is for the accommodation demands of M to be satisfied.

1.1.2.0. <u>A Naturalistic Aside and an Alternative Formulation</u>. The so called "model theoretic" arguments of Putnam (1978, 1980) have drawn attention to a general class of criticisms of causal theories of reference. According to those criticisms, causal theories of reference have the defect that there are <u>too many</u> different assignments of terms to referents which are compatible with the constraints on reference available to the causal theorist. In so far as some causal theories of reference face challenges along these lines, they do so because they conceive of natural kinds as phenomena (set theoretic or natural) which are otologically independent of the linguistic activity which underwrites reference to them. They then attempt to specify how the referential relation is established between these two sort of otologically unrelated phenomena.

The accommodationist proposal I offer here--which is, of course, a causal theory of reference <u>and</u> of kind definitions--avoids this pitfall; it <u>simultaneously</u> defines the reference relation, and the explanatory definitions and extensions of natural kind terms, in terms of the contributions which the actual deployment of those terms make to the achievement of accommodation between conceptual and methodological practices and relevant causal structures.

Another question about causal theories of reference concerns the relations which they posit between the reference relation and the referential (and other) intentions of members of the relevant linguistic communities. The accommodationist conception addresses this issue. Both the classificatory and predicative behavior mentioned in clause (1) and (more importantly) the accommodation demands mentioned in clause (2) are reflections of the intentions (referential and otherwise) of users of natural kind terms, and the accommodationist conception asserts that intentions play a role in establishing reference in just the ways indicated in these clauses.

This last point can be made clearer by a paraphrase of the conditions established by clauses (1) and (2) which--although it leaves out (without denying it) the important claim that natural kinds are ontologically dependent on linguistic practices--captures nicely the proposed connection between intentions and reference.

We can think of the programmatic definitions of the natural kinds referred to within a disciplinary matrix as characterizing those of the accommodation demands of practices within the matrix which can be satisfied, and thus as representing those of the intentions of the practitioners which, according to the accommodationist conception, are central in the establishment of reference. We may then paraphrase the accommodationist conception as maintaining that the explanatory definition (and thus the extension) of any natural kind is provided by an account of the family of properties shared by its members which underwrite the inductive/explanatory roles indicated by its true programmatic definitions with respect to the relevant disciplinary matrix.

1.1.2.1. Scorekeeping, II: Mind and Practice Dependence, Again. We've just made more precise the way in which--according to realist naturalism--natural kinds are mind and practice dependent.

1.1.3. Lawlessness. We have already seen that a certain sort of relativism is entailed by the accommodation thesis: the naturalness of a natural kind is relative to the accommodation demands of the disciplinary matrices within which reference to it contributes of successful explanation and induction. Realism and essentialism about natural kinds entail--given the actual causal complexity of the world--that natural kinds (or, at any rate, their naturalness) are, in that respect, practice (and, thus, mind) dependent.

We may further chip away at stereotypical conceptions of natural kinds by observing that the problem of accommodation to which natural kinds are (part of) the solution is in no way peculiar to the search for fundamental, exceptionless, eternal or ahistorical laws. Successful induction and explanation in the parochial, inexact, and historical sciences pose the same problems of projectibility as they do in the sciences (if there are any) which identify fundamental, exceptionless, eternal or ahistorical laws. Thus, in geology, economic history, evolutionary biology, and meteorology--where fundamental, exceptionless, eternal or ahistorical laws may be unavailable, even in principle--the kinds reference to which facilitates the satisfaction of accommodation demands will be every bit as natural natural kinds as those in the sciences more admired by logical positivists.

1.1.3.0. Score Keeping, III: No Laws. So now we have it that there can be natural kinds of which it is false that reference to them plays a role in stating "fundamental," "exceptionless," "eternal," or "ahistorical" laws. Indeed, we have seen that--at least for the purposes of a theory of projectibility and of natural kinds--the notion of a causally sustained regularity should replace that of a LAW, as that notion was (and is) understood in much of analytic philosophy. Nothing in the theory of projectibility, induction, natural kinds or reference suggests that there is any metaphysically important notion of a law beyond that of a causally sustained regularity. I suggest that there is none.

1.1.4. Theory Dependence and the Need for Truth. One important consequence of the accommodation thesis is that judgments about the definitions of scientifically significant natural kinds--and judgments about the projectibility of generalizations about them--are highly theory dependent. It could not be otherwise if the practices which involve deploying natural kind concepts and assessing the projectibility of generalizations are to be accommodated to esoteric causal structures in such a way that the epistemic access and accommodation conditions are satisfied.

In fact, hypotheses are properly judged as projectible just in case they are plausible in the light of the best established background theories (Boyd 1983, 1985a, 1985b, 1989, 1990, 1991, 1992). Judgments of projectibility--and thus the accommodation of inferential practices to causal structures--are dependent on background theories in two ways: Background theories determine the applications of the framework of natural kind terms within which the formulation of projectible hypotheses is possible, and background theories--often more specific ones--determine judgments of projectibility for hypotheses formulated within the relevant vocabulary.

The ways in which accommodation, both of classificatory practices and of inductive practices, are theory dependent adds another dimension to the relativity of naturalness to inferential architecture. In order for accommodation to be achieved, the background theories which determine how kind concepts are deployed in induction must be relevantly approximately true. Some approximation to the truth about the relevant kinds will be necessary in order that actual classificatory practices and other inferential practices are reliable enough to contribute to the satisfaction of the epistemic access condition.

More important is the fact that projectibility judgments regarding hypotheses formulated by reference to relevant natural kinds are theory dependent. It's not enough that an hypothesis can be formulated deploying suitable natural kind terms; recall that the hypothesis that all emeralds are grue is so formulable. Instead, hypotheses are assessed as projectible just in case they are plausible in the light of received theories. Here what is at issue is a sort of "indirect" or theory mediated evidential consideration: a theory counts as projectible just in case it is evidentially supported by a plausible inductive inference from well established background theories (see Boyd 1985b). Such inductive inferences achieve accommodation to causal structures, and thereby contribute to the satisfaction of the accommodation condition, only to the extent that the received background theories are relevantly true--only to the extent that they accurately reflect relevant causal structures. A kind is thus natural with respect to a theory dependent set of inductive practices if and only if the theories upon which those practices depend are relevantly approximately true. Accommodation requires truth (approximately).

1.1.5. Context of Invention, Accommodation and Another Need for Truth: There is another important dimension to the contribution which natural kinds, and the projectibility judgments their recognition facilitate, make to the establishment of reliable inductive practices. It lies in the theory of theory confirmation.

To a good first approximation, a theory receives significant evidential support from a body of successful predictions (or apparently successful explanations) just in case (a) the theory is itself projectible, (b) the observations or explaantions in question pit the theory's predictions (and/or its explanations) against those of its projectible rivals (for a closer approximation see Boyd 1982, 1985a).

In fact, the requirement that projectible theories be tested against their projectible rivals is a requirement for the rigorous testing of such theories. In testing a theory against its most plausible rivals we maximize the likelihood that it will be disconfirmed if it is in error. Projectibility judgments (that is, plausibility judgments) are matters of inductive inference from background theoretical premises to conclusions about the likelihood of various alternative answers to scientific questions. The fact that a proposed theory, or an alternative to it, is plausible in the light of previously confirmed theories is treated as significant, but not decisive, evidence for its (approximate) truth. Thus pitting all the projectible alternatives against each other experimentally or observationally is treated as a matter of choosing between all those alternative answers to a given scientific question for which there is already some significant positive evidence.

Given that these methods are the fundamentals of inductive inferences in science, we may ask, in the spirit of the accommodation thesis, what makes them inductively reliable? When and why are they accommodated to the relevant causal structures? Part of the answer, I suggest, is the obvious one: they are reliable only to the extent that the background theories in question approximately accurately represent the relevant causal structures and mechanisms. Only when this is true do projectibility judgments achieve accommodation between inferential practices and the world.

1.1.6. Accommodation, Social Architecture, and the "Context of Invention." Let us now ask, to what extent, and in what respects, must the background theories in question be approximately true and complete in order for the methods in question to function reliably? The answer, I suggest, establishes an especially intimate connection between the contribution of projectibility judgments to the reliability of scientific methods, on the one hand, and our capacities for theory invention on the other. The strategy for rigorous testing just described will prove generally reliable within a disciplinary matrix only when relevant background theories are accurate and complete enough that, for questions arising within the matrix, there is a significant tendency for one or more of the projectible answers actually invented and entertained as an answer to be relevantly close to the truth. When background theories are not sufficiently accurate and complete to ensure such a tendency, scientific methods are unreliable.

There is a further point here about the context of theory invention. The establishment of reliable theory dependent methods of the sort we are discussing--which is to say the establishment of accommodation--requires a certain theory mediated accommodation between relevant causal structures, the cognitive architecture of researchers, and the social architecture of their disciplines. It is not sufficient that accepted background theories be such as to, in fact, inductively support answers to scientific questions which (often enough) include good approximations to the truth. What is required is that the availability of relevant information, the structure of inductive practices and cognitive resources within the relevant scientific communities, and the social practices of communication and evaluation within them, be such that suitably articulated formulations of good approximations actually get formulated and that their predictive and explanatory resources be sufficiently widely appreciated. The lone, underfunded, socially marginalized researcher who makes the relevant inductive inference from background theories to the correct answer to a scientific question, but whose research does not get published in established journals, does not thereby make a contribution to the reliability of scientific practice.

Thus the accommodation of inferential practice upon which the epistemic reliability of scientific methods depends itself depends, not only upon the propositional structure of background theories, abstractly understood, but also upon historically contingent facts about the social and political structure of the relevant scientific communities and the broader society.

1.1.6.0. Keeping Score, IV: Historical Contingency and all that. Note that we now have it that the methods of science are socially, historically and politically situated constructs and that their epistemic reliability, such as it is, depends on (among other things) their social, historical and political situation. Thus we are well on our way to showing that the methods of science are, as relativist and anti-realist sounding slogans would have it, socially, historically and politically situated non-foundational methods.

2. Anti-Foundationalism.

2.0. Correspondence Truth. We'll return in Part 5 to consideration of the historical, social and political contingencies of scientific methodology when we discuss the political epistemology of knowledge of natural kinds and essences. For the present, what is important is that both knowledge about theoretical natural kinds, and the accommodation to causal structures which reference to them achieves, depend on the deployment of background theories which embody relevantly approximately true descriptions of causal phenomena. We turn now to the question of the nature of that truth.

2.0.0. The Metaphysical Innocence Thesis. When Kuhn (1970) describes the practice of "normal science" he makes points much like those rehearsed above. In doing normal science researchers investigating a scientific question know in advance, on the basis of knowledge embodied in the relevant paradigm, the (narrow) range within which they can anticipate that the answer will fall. This knowledge reflects a "quasi-metaphysical" understanding of the basic causal factors at work in the relevant disciplinary matrices. Scientists choose--in other words--between alternatives which are projectible by the lights of the received paradigm, AND IT IS BECAUSE THEIR CHOICES ARE THUS REGULATED--BECAUSE ACCOMMODATION IS THUS ACHIEVED--THAT PARADIGM GOVERNED RESEARCH IS SUCCESSFUL, IN SO FAR AS IT IS SUCCESSFUL.

Kuhn is often counted as one of the founders of contemporary neo-Kantian philosophy of science, despite his claims to be a kind of scientific realist. According to the neo-Kantian view attributed to him (which he certainly seems to accept; see Kuhn 1970, pp 101-102) fundamental laws within a paradigm can be thought of as, in some important sense, truths by convention. Since these are among the components of a paradigm which determine projectibility judgments, the Kuhnian view entails that truths by convention can provide (some of) the theoretical knowledge necessary to effect the success of paradigm governed research.

Can this be so? Can conventional knowledge play this role? In the light of the accommodation thesis, the question becomes this one: can claims about causal relations be true by convention; can we make causal claims true or false by adopting rules of language, or conceptual schemes, or paradigms, or the like?

By the metaphysical innocence thesis I understand the common sense view that human social practices are metaphysically innocent: that we cannot make metaphysical propositions (in this particular case, causal propositions) true or false simply be adopting particular conventions or practices. Of course it is not in doubt that our linguistic conventions and other social practices contribute to establishing the truth conditions for, for example, causal claims. Obviously the semantics of causal discourse is established in part by our conventions and practices. What is at issue is whether or not social or linguistic practices can establish causal (or other metaphysical) facts or states of affairs.

The answer provided by the metaphysical innocence thesis ("no") is just the answer which is ordinarily simply presupposed whenever philosophers argue that some apparently metaphysical issue is better thought of instead as settled by matters of linguistic convention, or whenever scientists treat some choice between different conventions for representing natural phenomena as a matter of pedagogical or cognitive or computational convenience rather than a matter of substance. If the metaphysical innocence thesis can be sustained, then real (correspondence) truth and approximate truth will be required for accommodation, not neo-Kantian "socially constructed" truth.

2.0.1. The No Non-Causal Contribution Thesis, I: Historical Evidence. If we explore this issue further we see that, strictly speaking, it isn't true--and it isn't presupposed by the philosophical and scientific methods just discussed--that the adoption of linguistic conventions, social practices, and the like, make no contribution to establishing causal states of affairs. Of course they do. Social activities which underwrite and sustain linguistic conventions are themselves causal phenomena, so they have a causal impact on other causal relations and properties. The version of the metaphysical innocence thesis which is presupposed by the methods in question is the no non-causal contribution thesis (henceforth: 2N2C): the thesis that human social practice make no non-causal contribution to causal properties and relations. What 2N2C denies is that there is some further sort of contribution (logical, conceptual, socially constructive, or the like), explicable by distinctly philosophical rather than empirical theories, which the adoption of theories, conceptual frameworks, and the like makes to the establishment of causal powers and relations.

Of course, the fact that 2N2C nicely underwrites some highly plausible philosophical and scientific methods does not guarantee that it is true. I have argued for it in some detail elsewhere (Boyd 1990, 1991, 1992). For the present, I'll content myself with offering two considerations in its favor.

In the first place, as early critics of Kuhnian social constructivism remarked, the phenomenon of anomalies within paradigm guided research--especially anomalies recognized by researchers operating within the paradigm--seems to constitute an empirical refutation of the sort of determination of causal phenomena by social practices required by neo-Kantian conceptions. In a world in which causal relations conformed to Newtonian mechanics (even if that conformity was established by social convention), there would be no non-Newtonian anomalous phenomena. Kuhn (rightly) insists that there are such anomalies and that scientists operating within the Newtonian paradigm detected them. Thus the neo-Kantian conception of causal relations (which, in the light of the accommodation thesis, the view attributed to Kuhn entails) appears to be refuted.

Of course it is possible to reply that Kuhnian paradigms are not the counterexamples to 2N2C; perhaps there are others. The prospects for this reply are, however, not promising. In the first place, Kuhn's account of paradigm governed science and other indications of the deep theory dependence of scientific methods provide the best arguments thus far for neo-Kantian conceptions of science, so if 2N2C holds true for those cases it seems pretty secure.

A broader point illustrated by the example of scientific paradigms is that neo-Kantian conceptions of causation, unlike Kantian ones, posits variability, from one paradigm to another, in the causal phenomena participants encounter. It thus predicts that we should observe a sort of wishing-makes-it-so phenomenon, with participants in different paradigms (together with their instruments) responding to different and incompatible causal influences even under (what we would ordinarily take to be) the same circumstances. It is the fact that the historical evidence seems to point towards the opposite estimate of the effects of wishing--especially in the sciences--which counts heavily against neo-Kantian conceptions.

2.0.2. The No Non-Causal Contribution Thesis, II :Conventions, Arbitrariness, and Methodology. There is a related argument in favor of 2N2C which I find even more compelling, perhaps because it relies on considerations from absolutely ordinary everyday science without even a hint that "revolutionary" scientific practice is at issue. I contend that scientific methodology regarding conventionality in description and measurement fundamentally presupposes 2N2C. Scientists count conventional features of their descriptive apparatus as methodologically irrelevant in ways which would not be reliable if 2N2C were mistaken.

What I have in mind involves the acceptance by scientists in practice (in theory, too, for that matter) of the principle of methodological equifertility of conventionally established representations (Boyd 1990, 1991, 1992). According to the equifertility principle, when a body of theories or other representations appears to license a methodological practice, the license is epistemically cogent only if the same practice is licensed by every other representation such that the choice between it and the representation in question is purely conventional. Methodologically speaking, conventions don't count. It is this principle which, for example, would rule out the methodology of an study of the political economy of international trade whose conclusions would be different depending on whether the input data on prices of commodities were expressed in dollars or German marks, or of a study of physical phenomena whose results depend on whether length was measured in centimeters or inches.

This principle is ubiquitous in its application in everyday as well as extraordinary science. It is as close to obvious as methodological principles get. Still, it's general reliability seems to depend on 2N2C. If there are cases in which relevant causal structures are non-causally determined, in some neo-Kantian way, by the representational schemes scientists adopt, then these will be cases in which (some applications of) the equifertility principle will fail. Perhaps this is so; perhaps 2N2C holds just for non-world-making conventional choices. Or perhaps 2N2C holds for the cases philosophers have thus far considered but fails for sorts of cases unanticipated either by neo-Kantians or their critics. Perhaps, but the methodological centrality of the equifertility principle, and the utter ubiquity of its (apparently successful) applications, place a strong burden of proof on the defenders of neo-Kantian conceptions.

2.0.3. Anti-(neo)-Kantian Conclusion. I conclude that consideration of the available scientific evidence and of the foundations of apparently reliable central scientific methods points to the conclusion that 2N2C obtains. We do not observe the paradigm-or-culture relativity of causal phenomena apparently predicted by neo-Kantian conceptions, and powerful, central and ubiquitously applicable scientific methods presuppose its falsity with respect to every case in which they are applied. It is difficult to escape the conclusion that, while we and the world jointly determine the definitions of kinds, we do not have a (non-causal) hand in making causal facts. Our role in making causal facts is merely the metaphysically modest causal role which we play because we are causal phenomena ourselves.

2.1. Correspondence Truth and the Failure of Foundationalism. Theory-dependent methods are reliable only in so far as the background theories which determine their dictates are approximately true descriptions of causal phenomena. We have just seen that neo-Kantian conventionalist conceptions of the relevant sort of truth is inadequate. Thus, as promised, we see that a realist conception of correspondence truth (and of approximation) is required for an understanding of how the accommodation demands of disciplinary matrices are met.

Let's turn to the issue of foundationalism. By a foundationalist conception of knowledge we ordinarily understand a conception according to which all beliefs which are cases of knowledge are ultimately justifiable from a class of epistemically privileged foundational beliefs *via* inference rules which are themselves justifiable either *a priori* or on the basis of foundational beliefs. What we have just seen is that the inductive methods in a particular science at a given time are epistemically reliable only if

a. the conceptual and classificatory resources they deploy are suitably accommodated to relevant causal structures, and

b. the background theories which determine their dictates provide relevantly approximately true representations of those causal structures, and

c. the social and political structures of the relevant scientific communities, and their cultural environments tends systematically to lead to epistemically appropriate patterns of theory invention.

If causal conditions like these could be made true by convention--if wishing could make it so--then we might have *a priori* knowledge that a.-c. obtain in particular cases. But 2N2C is true, so wishing doesn't make it so. Plainly, for no actual case of inductive inference in the sciences will a.-c. be themselves foundational.

Thus the inductive methods of the sciences are non-foundational: their epistemic reliability, when they are reliable, depends upon quite contingent features of the intellectual, social and political history of the relevant communities. In many scientific disciplines in recent times methods have been, with respect to a wide range of issues, spectacularly reliable, but the depth of that reliability, and of our confidence in it, does nothing to show that those methods are foundational. They aren't.

2.1.0. Keeping Score, V: Historical Contingency and all that, Again. OK, now we have it. The methods of science are, as relativist and anti-realist sounding slogans would have it, socially, historically and politically situated non-foundational methods.

3. Bicameralism.

3.0. Locke, Again: The Metaphysics of Natural Kinds. Locke maintained that while Nature makes things similar and different, kinds are "the workmanship of men." The lesson we should draw from the accommodation thesis 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[2]. A natural kind just is the implementation, in language and in conceptual, experimental and inferential practice, of a (component of) a way of satisfying the accommodation demands of a disciplinary matrix. Natural kinds are features, not of the world outside our practice, but of the ways in which that practice engages with the rest of the world. Biological taxonomists sometimes speak of the "erection" of biological taxa, treating such taxa as, in a sense, human constructions. They are right--and the same thing is true of natural kinds in general.

Locke said that "...each abstract idea, with a name to it, makes a distinct Species." His conception was that kinds are established by a sort of unicameral linguistic legislation: people get to establish kind definitions by whatever conventions (nominal essences) for the use of general terms they choose to adopt.

According to the accommodation thesis, natural kinds are, instead products of bicameral legislation in which the (causal structure of the) world plays a heavy legislative role. 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 (remember that natural kinds are legislative achievements--that is, artifacts) 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. Natural kinds are the workmanship of women and men.

The causal structures in the world to which accommodation is required are, of course, independent of our practices in the sense specified by 2N2C. Still, natural kinds are social artifacts. No natural kinds exist independently of practice. The kind natural kind is itself a natural kind in the theory of our inferential practice. That's why the reality of kinds needs to be understood in terms of the satisfaction of the accommodation demands of the relevant disciplinary matrices.

---

2. 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.

3.0.0. Keeping Score, VI: Natural Kinds are Social Constructions. We've just seen that this is true with respect to the philosophically central notion of natural kinds. Of course, we can talk about, say, undiscovered natural kinds, but this is best understood as talk about possible extensions of the conceptual resources of our disciplinary matrices to achieve further accommodation to relevant causal structures. Similarly, we may follow Quine (1970) in extending the notion of a natural kind to encompass the accommodation-related conceptual resources of other creatures. Even so, "postmodernists" and other relativists are right to think of paradigm natural kinds as human social constructions.

3.1. Property-Clusters and Natural Kinds. Biological species and chemical elements and compounds are the paradigmatic philosophical examples of natural kinds. I have argued elsewhere (Boyd 1988, 1989, 1991, 1993, forthcoming b) that there are a number of scientifically important natural kinds (properties, relations, etc.), biological species among them, whose natural definitions are very much like the property-cluster definitions postulated by ordinary-language philosophers except that the unity of the properties in the defining cluster is mainly causal rather than conceptual.

The natural explanatory definition of one of these homeostatic property cluster kinds is provided by the members of a cluster of often co-occurring properties and by the ("homeostatic") mechanisms that bring about their co-occurrence. It is an a posteriori theoretical question which of these properties and which of the homeostatic mechanism count, and to what extent they count, in determining the explanatory definition (and, thus, the membership conditions) for the kind. In cases of imperfect homeostasis in which some of the properties in the cluster are absent, or some of the mechanisms inoperative, it will sometimes happen that neither theoretical nor methodological considerations assign the object being classified determinately to the kind or to its complement, with the result that the homeostatic property cluster definition fails to specify necessary and sufficient conditions for kind membership. Both the property-cluster form of such definitions and the associated indeterminacy are dictated by the fundamental epistemic task of accommodating inferential structures in the relevant disciplinary matrices to relevant causal structures. In particular, the indeterminacy in extension of these natural definitions could not be remedied without rendering the definitions un-natural in the sense dictated by the accommodation thesis.

The paradigm cases of natural kinds--biological species--are homeostatic cluster kinds. The appropriateness of any particular biological species for induction and explanation in biology depends upon the imperfectly shared and homeostatically related morphological, physiological and behavioral features which characterize its members. The definitional role of mechanisms of homeostasis is reflected in, for example, the role which interbreeding between conspecific populations, and reproductive isolation from contraspecific ones, plays in defining some sexually reproducing species. For some sexually reproducing species, the exchange of genetic material between their populations, and the absence of such exchange with other related species, is essential to the homeostatic unity of the other properties characteristic of the species.

The necessary indeterminacy in extension of species terms is a consequence of evolutionary theory, as Darwin observed: speciation depends on the existence of populations which are intermediate between the parent species and the emerging one. Any "refinement" of classification which artificially eliminated the resulting indeterminacy in classification would obscure the central fact about heritable variations in phenotype upon which biological evolution depends, and would thus undermine the accommodation of the classificatory resources of biology to relevant causal structures.

Biological species also exhibit another important characteristic which they share with many other HPC natural kinds. The homeostatic property cluster which serves to define a biological species is not individuated extensionally. Instead, the property cluster is individuated like an historical object or process: certain changes over time (or in space) in the property cluster, or in the underlying homeostatic mechanisms, preserve the identity of the defining cluster. In consequence, the properties which determine the explanatory definition of a species (and, thus, the conditions for membership in it) may vary over time (or space), <u>while</u> <u>it</u> <u>continues</u> <u>to</u> <u>have</u> <u>numerically</u> <u>the</u> <u>same</u> <u>definition</u>. The historicity of the individuation criterion for the definitional property cluster reflects the explanatory or inductive significance of biological species (the programmatic definitions of particular species and of the category <u>species</u>). A species is a manifestation of a persisting but local homeostatic stability in evolutionary processes. The properties of its members, and the homeostatic mechanisms which, together, underwrite the stability of a species (and thus constitute its explanatory definition) can vary over time and space.

This is clearest in the case of a sexually reproducing species whose integrity over time involves reproductive isolation from other closely related species. At any given time, whatever properties of the members of such a species explain its reproductive isolation from related species will be part of its explanatory definition. As some closely related species become extinct, or others emerge, the integrity of the species may depend in this way on different properties of their members at different times. So, its defining cluster can consist of different properties and mechanisms at different times.

Another important property of the explanatory definitions of biological species is also illustrated by cases in which their integrity is ensured by reproductive isolation from related species. Often such isolation depends on environmental conditions--for example conditions in which breeding behavior in two closely related species, which would otherwise interbreed, occurs at different times of the year. Thus among the properties which make up the explanatory definition of a biological species there can be some which are relational rather than intrinsic.

Relational properties figure in the explanatory definitions of biological species in another way. Among the evolutionarily important properties which members of a biological species share are those which ensure that populations of the species exhibit similar evolutionary tendencies--that's why species-level classification is so important in achieving accommodation in evolutionary biology. It is widely agreed among biologists and philosophers of biology that, for this reason, species are <u>historically</u> delimited. If we suppose that, in some contemporary pond, there happens to be a population of organisms physically identical in every respect to a population of some early Jurassic fish, they would not be members of the same species. The reason, it is agreed, is that external relational factors (which will be profoundly different between Jurassic ponds and contemporary ponds) are so crucial to the evolutionary fate of a species that the contemporary organisms and their Jurassic analogues would not share the same "evolutionary fate" to the extent appropriate for conspecificity. The explanatory definition of a biological species has an irreducible historical (and historically delimited) character.

3.1.0. Keeping Score, VII: Historicality and all that. According to one version of realist naturalism, the definition of a natural kind is always provided by a set of eternal, unchanging, ahistorical, and intrinsic necessary and sufficient conditions. We have just seen that a consistently developed realist naturalism entails that the explanatory definitions of some paradigm natural kinds--biological species--have none of these properties.

3.2. The Final Score. We began by considering a stereotype of the realist conception of natural kinds according to which they are

1. independent of human practices;

2. defined by

    a. eternal,

    b. unchanging,

    c. ahistorical, and

    d. intrinsic

necessary and sufficient membership conditions;

3. referred to in

    a. fundamental,

    b. exceptionless,

    c. eternal, and

    d. ahistorical

laws; and

4. discovered by the deployment of

    a. eternal,

    b. ahistorical, and

c. foundational

scientific methods.

Relativist and anti-realist sounding "anti-essentialist" slogans popular outside mainstream analytic philosophy suggest that in every respect this stereotypical conception is mistaken. What we have just seen is that consistently developed <u>realist</u> <u>naturalism</u> about natural kinds and their essences entails that these slogans are largely correct. Natural kinds are:

1. <u>always</u> interest and practice dependent social constructions; and

2. <u>often</u> defined by

    a. open textured (neither necessary nor sufficient),

    b. historically situated,

    c. relationally and historically defined, (ant thus)

    d. non-eternal, and

    e. non-intrinsic

properties.

We have seen that many of them are referred to in approximate and historically specific causally sustained generalization but not in any laws which are

3. a. fundamental,

    b. exceptionless,

    c. eternal, or

    d. ahistorical,

and that <u>all</u> of them are studied by scientific methods which are

4. a. socially situated,

    b. historically situated,

    c. politically situated, and

    d. non-foundational.

The score (by my count) is

Realist stereotypes: 0

Relativist slogans: 14.

[But remember, truth is correspondence truth, our social practices are metaphysically innocent, natural kinds possess a posteriori definitions and knowledge is a matter of causally reliable methods. Four cheers for realism.]

4. Water, H2O, and all that: Natural Kinds, Modality, and the Limits of Linguistic Legislation.

4.0. The Problem. I suggested earlier that a properly developed realist naturalism about the essences of natural kinds would involve a critique of the current methodological enthusiasm for exploring the properties which natural kinds have "in all possible worlds."

My interest in this topic arose from the suspicion that there was something fundamentally wrong with the practice of testing philosophical theories--in ethics, say, or epistemology--by comparing their dictates with the deliverances of our philosophical intuitions about which actions are good, or which beliefs justified, in possible worlds so different from the actual world that a great many of the regularities we ordinarily rely on in moral or epistemic judgments fail to obtain. I wanted to understand as a matter of philosophical theory what was wrong with relying on such intuitions. What I concluded was that in many (probably all) such cases there is no fact of the matter about what would be good or bad, justified or not justified under such bizarre conditions. I here recruit the accommodation thesis, and the conception of the modal properties of natural kinds which it underwrites, to defend that conclusion.

4.1. The Limits of Linguistic Legislation: The Basic Idea. According to the accommodation thesis, natural kinds are products of bicameral linguistic legislation; their defining conditions are determined (only) by the ways in which linguistic practices contribute to the satisfaction of accommodation demands of disciplinary matrices. In particular, the membership conditions for natural kinds in other possible worlds is determined by our linguistic legislation in the actual world. What I'll argue is that the bicameral legislative authority which we and the (actual) world possess does not extend to establishing such membership conditions for sufficiently weird possible worlds.

What I will argue is that there are three distinct ways in which the authority of that legislation regarding non-actual worlds is limited. I'll begin with the case of homeostatic property cluster kinds (relations, etc.)--which include, I believe, all or most of the natural kinds and relations (like, e.g., reference, knowledge, justification, moral goodness) central to the disciplinary matrix to which philosophy itself belongs (see Boyd 1988, 1993).

4.2. The Limits of Linguistic Legislation, I: Homeostatic Property Cluster Kinds. In the case of an HPC kind, an explanatory definition is provided by a (perhaps historically individuated) process of homeostatic property clustering in the actual world. An explanatory definition of a biological species, for example, is provided by a family of approximately shared features together with the clustering mechanisms which sustain their homeostatic unity.

In a non-actual possible world an object is in such a kind just in case it relevantly participates in the very same process in that world. Thus, for example, if we inquire about whether or not raccoons would adapt or go extinct if near arctic conditions were to slowly develop throughout their range, we are inquiring about possible worlds (corresponding to non-actual possible futures for the actual world) in which weather changes throughout the range of Procyon lotor, and the homeostatic property clustering characteristic of that species either changes in response (so that somewhat different properties and mechanisms lie within it) or ceases to exist because the members of the species all die out. We are thus inquiring about different possible future histories of the very same (non-extensionally individuated) property clustering which constitutes, in the actual world, the definition of the kind Procyon lotor. Such a definition is thus an historical process like, e.g., the second world war or the emergence of capitalism.

The same is true for many HPC definitions: they are process-like historical phenomena (token or type depending on the kind in question). They are like wars and economic transitions (and like persons) in that (numerically) very same HPC definition can have different constituents in different possible worlds. They are also like persons and historical processes in that, for sufficiently distant possible worlds--in which some but not many of their actual world constituents are manifested together--it becomes indeterminate whether or not they exist. Thus, for P. lotor, as for any other HPC kind, as one moves away from the actual world there is an (indeterminate) zone of possible worlds where it is indeterminate whether it's definition is manifested at all. Such worlds lie beyond the authority of the linguistic legislation which establishes the HPC kind in question (for a more extensive discussion see Boyd 1988).

4.3. Linguistic Legislation and Disciplinary Matrices as HPC Phenomena.

4.3.0. Kinds and Their Disciplinary Matrices in Non-Actual Worlds. I propose to extend the arguments just presented to indicate one way in which our bicameral linguistic legislation lacks authority regarding distant possible worlds, whether the kinds in question are HPC kinds or not.

Recall that the explanatory definitions of a kind referred to within a disciplinary matrix are to be understood in terms of the properties required of its members for the satisfaction of accommodation demands of the matrix in question; natural kinds get their (actual world) explanatory definitions (and, thus, their membership conditions) from the ways in which accommodation is actually accomplished. What I now propose is that, in a similar way, the explanatory definitions of an actual world natural kind in a non-actual possible world is determined by the conditions for the satisfaction of accommodation demands in that world. Here's the conception I favor:

Let M be an actual world disciplinary matrix, $t_1,...t_n$ the natural kind terms deployed within the discourse central to the inductive/explanatory successes of M, $k_1,...,k_n$ the kinds they refer to in the actual world, and W a possible world. Let ^{W}M be the classificatory, inductive and explanatory practices of M as they would have to be implemented when **exported** for application **regarding** W.

Then, $k_1,...,k_n$ are well defined in W, and the families ${}^{W}F_1,...,{}^{W}F_n$ provide their explanatory definitions in W and determine their extensions in W, just in case:

1. (Epistemic access condition) There is a systematic, tendency, causally sustained by the relations between ${}^{W}M$ and causal structures in W, for what is predicated of t_i within the practice of the **exported** matrix ${}^{W}M$ to be approximately true of things which in W satisfy ${}^{W}F_i$, i=1,...n.

2. (Accommodation condition) This fact, together with the causal powers in W of things satisfying these explanatory definitions, causally explains how the use of $t_1,...t_n$ in ${}^{W}M$ contributes to accommodation of the **exported** inferential practices of ${}^{W}M$ to relevant causal structures in W: that is to the tendency for **the exported practice** ${}^{W}M$ to identify causally sustained generalizations in W and to obtain correct explanations for phenomena in W.

Except for the deployment of the notion of **the exportation** of a disciplinary matrix for application in a possible world, this is just the relativisation to W of the accommodationist conception of kind definitions offered in section 1.1.2. Like that conception it can be paraphrased in terms of the relation between programmatic and explanatory definitions. The explanatory definition of a natural kind in a possible world, W, is provided by an account of the family of properties shared by its members in W which underwrite the inductive/explanatory roles indicated by its programmatic definition for W, with respect to the relevant **exported** disciplinary matrix.

So, an object x will, in W, be in k_i (i=1,...,n) in W just in case so classifying it is central to its being the case that causal relations among things so classified in W will result in the satisfaction of the accommodation demands of ${}^{W}M$ as implemented for W. ${}^{W}F_i$ will differ from F_i in just whatever ways are required to preserve the sort of accommodation achieved in the actual world through the use of t_i in M, given the ways in which W differs in its causal structure from the actual world.

The alternative formulation in terms of the relation between programmatic and explanatory definitions allows us to make precise two different ways in which the explanatory definition of a natural kind might come to be different in some non-actual world. In the simplest case, the possible world in question might be sufficiently like the actual world in relevant respects that exactly the same disciplinarily appropriate inductive and explanatory projects are appropriate to the relevant disciplinary matrix in both worlds. In such a case, differences in membership conditions for natural kinds between the two worlds will be entirely a matter of the different properties which are causally relevant, in the two worlds, for the satisfaction of shared programmatic definitions of the relevant kinds.

In more complex cases, even the programmatic definitions of the relevant kinds may be somewhat different. Recall that, on the conception defended here, the programmatic definitions of natural kinds reflect the successful explanatory and inductive projects characteristic of the relevant disciplinary matrices. In a non-actual world, W, some of the inductive and explanatory projects characteristic of an actual world matrix may not be achievable; in the programmatic definitions of the relevant kinds for the corresponding matrix for that world, reference to those projects may be replaced by reference to other analogous disciplinary projects, unrealizable in the actual world but realizable in W. In such cases, explanatory definitions of kinds may differ from world to world at least partly because of these difference in their programmatic definitions.

It remains to explain the relevant notion of the exportation of an actual world disciplinary matrix for implementation regarding a non-actual world. Here's the basic idea: actual world membership in a particular actual world natural kind is a complex causal capacity defined with respect to the accommodation demands of an actual-world disciplinary matrix; to be a member of such a kind in some non-actual world is to have the right causal capacities with respect to the accommodation demands of the very same disciplinary matrix, understood as an explanatory and inductive project implemented in that other world. This conception seems to answer to our standard philosophical practice: asked what the definition of, say, a biological natural kind would be in some possible world, we imagine ourselves doing biology in that world and ask what definition would be appropriate.

In the account I offer, instead of maintaining that the explanatory definitions of kinds in a non-actual possible world
are fixed by the accommodation demands of the relevant disciplinary matrix in that world, I speak instead of the matrix in question being exported for implementation regarding that world. We often evaluate claims about natural kinds in possible worlds in which there could not be any disciplinary practices at all. For example we know that the natural kinds of chemistry would have the same definitions in any possible world just like the actual one except that it contains no cognizing systems--and thus no disciplinary practices. Moreover, we know that those kinds would have the same definitions in some possible world in which there were no places which were not so toxic that no living practitioners could implement the relevant disciplinary matrix. In such cases what we do in practice is to envision implementing the inductive and explanatory aims of the relevant disciplinary matrix from a position which somehow affords us a birds-eye view of the non-actual world without actually being in it.

Exactly how this abstraction is to be understood is hairy. One approach--which suffers from an additional epicycle of complexity--is to think of the implementation of a matrix, M, for a possible world, W, in terms of the implementation of M in some closely related possible world W', which is just like W except that, in W', somehow or other the practices of M are implementable (perhaps as a result of local exceptions to causal regularities which obtain in W) and get implemented. Another approach--which suffers the disadvantage that it makes it harder to account for the role of instrumentation in relevant disciplinary matrices--is to think of the matrix M as somehow embodying an abstract set of classificatory and inferential procedures whose determinations regarding a world, W, can be identified independently of any conception of their social or technical realization in W or in any closely related world.

Since I am inclined to view "possible worlds" as somewhat vague conceptual devices deployed by philosophers to (sometimes helpfully) regiment modal discourse, rather than as determinate entities of some sort, I suspect that the details of an explication of the notion of the implementation of a matrix for a world are not especially important. Readers with more sensitive views about the ontological status of possible worlds are invited to develop their own analyses of our philosophical practice of considering the implementation of matrices for possible worlds in determining the extensions of natural kind terms in them. What is important here is that that is what we do.

4.3.1. Why this is what we Should Do. Let's call the practice just discussed--of determining the extension of natural kind terms in other possible worlds by (somehow or other) exporting to those worlds the relevant disciplinary practices--the exportation strategy; by the exportation thesis let's understand the thesis that the exportation strategy is the right one to deploy in these matters.

The exportation thesis nicely rationalizes the details of the treatment offered here of HPC kinds; in consequence, it also rationalizes counterfactuals about membership conditions for biological species which are independently plausible. Nevertheless, it has consequences which might seem to cast doubts on its appropriateness. For example it has the consequence that the properties definitive of a natural kind can vary across possible worlds. One effect of the prominence of examples like "Water=H20" has been to suggest just the opposite conclusion. Thus it is a reasonable question whether or not the exportation thesis is true.

4.3.2. Contrastive Actualism. I'll briefly sketch an argument in favor of the exportation thesis. It rests on a thesis (contrastive actualism) about the semantics of counterfactual statements about actual world causal phenomena--statements about the properties which such phenomena have in non-actual worlds. According to contrastive actualism, statements, made in the actual world, about what happens to actual world causal phenomena in non-actual possible worlds are true or false depending on whether or not they accurately reflect, via what we might call contrastive specification, facts about causal powers of, and causal relations between, relevant phenomena in the actual world. By contrastive specification I have in mind a technique for specifying alleged causal powers of, or causal relations between, actual world phenomena by describing the causal powers of or causal relations between these, or relevantly similar, phenomena in possible worlds which contrast with the actual world in specified ways.

Some examples will illustrate the way I take contrastive specification of claims about actual world causal phenomena to work. When someone says that there would be no mammals in a possible world in which there was no atmospheric oxygen, what she says is true just because of the central causal role oxygen plays in sustaining mammalian life in the actual world. If she goes on to say that there would be no mammals even in possible worlds with atmospheric oxygen if oxygen molecules were larger than some specified size, a, what she said is true if, for example, in the actual world, the permeability of mammalian lung tissues to oxygen depends on oxygen molecules being smaller than a.

One last illustration will indicate more fully how contrastive specification works in cases more closely analogous to those concerning the explanatory definitions (and, thus, the extensions) of natural kind terms in non-actual worlds. Giant pandas, I understand, eat only varieties of bamboo. This makes true the claim that in possible worlds just like the actual world except that in them all bamboo species become extinct, pandas also become extinct. Suppose that a biologist, having explained this to us, goes on claim that in another possible world--one just like the actual world except that, instead of bamboo, it had some other logically possible plant species, of some sort, S, which she describes in detail--pandas would survive by eating plants of those species.

What would make her statement true? The answer, I propose, is that it would be true just in case the properties she attributes to the logically possible plants in the non-actual world in question include among them all of the properties of actual world bamboo plants which suits them (and no other actual world plants) to play the nutritional role they do as food for giant pandas and if the behavioral dispositions and genetic makeup of actual world giant pandas is such that either the causal properties of the possible plants in question would trigger feeding behavior in pandas generally or those properties would trigger feeding behavior in some sufficient number of (perhaps genetically atypical) pandas, so that natural selection would lead to the evolution of new feeding behaviors before giant pandas became extinct.

In this case, as in the others we have considered, the truth or falsity of a statement made in the actual world about causal phenomena in a non-actual world was a matter of the way in which that statement succeeded or failed in accurately contrastively specifying the causal properties of phenomena in the actual world. The same is true, according to contrastive actualism, for all statements made in the actual world about the properties of actual world causal phenomena in non-actual worlds.

I will not defend contrastive actualism here beyond indicating that it seems to capture a key semantic role of counterfactuals, and remarking that even someone who holds that the truth of causal statements about actual world phenomena is ultimately eliminatively reducible to the truth of related counterfactuals can still accept contrastive actualism as a constraint on her conception of the relevant eliminative reduction.

4.3.3. Defending Exportation. In order to understand the relevance of contrastive actualism to the defense of the exportation thesis, we need to return once more to giant pandas. Recall the counterfactual claim we considered to the effect that giant pandas survive in possible worlds containing a particular sort of logically possible but non-actual plant. Counterfactual statements of this sort, let's call them sustenance counterfactuals, when made in the actual world, express claims about how persisting phenomena in the actual world are sustained in the actual world by indicating how they would or would not be sustained in non-actual possible worlds.

We are now in a position to see why the exportation thesis is true. Actual world membership conditions for natural kinds--their explanatory definitions--are a matter of what causal factors explain how the satisfaction of the accommodation demands of the relevant disciplinary matrices are actually achieved. This is precisely a matter of the sustenance of the explanatory and inductive successes of those matrices. So counterfactuals of the form "things with properties P_1,...P_n would be in K in W," for P_1,...P_n properties, K a natural kind, and W a non-actual possible world, are sustenance counterfactuals: they make, via contrastive specification, claims about the ways in which accommodation is sustained in actual world disciplinary matrices by indicating how the accommodation demands of those matrices would or would not be sustained in other possible worlds. This is just what the exportation requires: that the membership conditions for natural kinds in a non-actual world reflect exactly those changes in classificatory practice which would be required to sustain the accommodation of the relevant disciplinary matrix implemented for the particular (contrasting) conditions in the non-actual world in question.

4.3.4. <u>The Limits of Linguistic Legislation, Again</u>. It remains to see why the exportation thesis entails an additional limitation of the authority of our bicameral linguistic legislation--one which operates even when the natural kinds in question are not themselves HPC kids. Once again, giant pandas can help us out. Consider counterfactual statements of the form "Giant pandas survive in W," for various non-actual worlds W. Giant pandas form a biological species, an HPC kind. Therefore, as we turn out attention to possible words sufficiently different from the actual world in relevant ways, the identification of giant pandas in those worlds becomes indeterminate. Our bicameral legislation does not extend to the determination of truth values for sustenance counterfactuals about pandas in such worlds.

In general, our bicameral legislation determines truth values for sustenance counterfactuals about a phenomenon P only for those possible worlds where it determines conditions for <u>being P</u>. When P is itself an HPC phenomenon, the limitations we have already recognized for the authority of linguistic legislation with respect to HPC kinds will limit our legislative authority with respect to sustenance counterfactuals about P.

What I now want to indicate is that similar considerations show that our legislative authority is limited even with respect to non-HPC natural kinds. We have already seen that a crude but illustrative way of articulating the exportation thesis is this: one must classify biological organisms in a possible world, W, just as <u>biology done in W</u> would require; and similarly for physical kinds and physics, for chemical kinds and chemistry, etc.

Despite its crudeness, this formulation allows us to see an important point about the definitions of natural kinds in non-actual worlds. On the simplified formulation the extension of a biological kind, k, is determined in a non-actual world, W, <u>only if</u> there is a sufficiently determinate answer to the question, "What is the manifestation <u>in W</u> of biology?" When we move from this simplified formulation to one which speaks of disciplinary matrices for a world, W rather than disciplines in W, the same condition applies: the extensions of the family of natural kind terms characteristic of a disciplinary matrix, M, is determined in a possible world, W, <u>only if</u> there is a sufficiently determinate answer to the question "What is the manifestation <u>for W</u> of the actual world disciplinary matrix M?"

Actual world disciplinary matrices are families of social and instrumental practices which are themselves HPC phenomena. What characterizes disciplinary matrices is that practice within such a matrix involves the simultaneous satisfaction of a very large number and wide range of accommodation demands. It is characteristic of natural kind terms that, when their use us suitably accommodated to causal structures, the satisfaction of some of the accommodation demands of a matrix is conductive to the satisfaction of the others. This <u>homeostatic unity</u> of instances of accommodation demand satisfaction is what defines a discipline or disciplinary matrix <u>and</u> what defines the subject matter of the philosophical theory of natural kinds.

Thus, the individuation of disciplinary matrices across possible worlds poses the same problems which face the individuation of any other HPC phenomena, like species, across possible worlds. For possible worlds sufficiently distant from the actual world it becomes indeterminate what the manifestation is, for that world, of any given actual world disciplinary matrix, or even whether there could be any such manifestation. In such cases the explanatory definitions of the relevant natural kinds will likewise be indeterminate. Claims about the relevant natural kinds in such worlds will lack truth values. In so far as natural kinds and kind terms go, we (and the actual world) are simply not capable of legislating "for all possible worlds."

4.4. Constitutive Description: One More Limit to the Authority of Linguistic Legislation.

4.4.0. Constitutive Description. There is (at least) one more way in which the authority of our linguistic legislation is limited with respect to the modal properties of natural kinds. It has to do with the implementation of a actual world disciplinary matrix, M, for a non-actual world, W, where M is well defined in or for W but where its implementation for W would require a denotational refinement (in the sense of Field 1973) in the use of one or more natural kind terms.

What I have in mind is analogous to certain actual world cases of indeterminate linguistic legislation. Consider the use of the term "element," and of the terms for the various chemical elements, prior to the discovery of isotopes. The inductive and explanatory practices of the disciplinary matrix which included chemistry anticipated that all samples of substances characterized by the same location in the periodic table of the elements would share all the same physical and chemical properties, and this expectation was reflected in how the term "element" and the names of particular elements, like "hydrogen" and "carbon," were deployed--that is, in the prevailing conceptions of their programmatic definitions. What we now describe as the discovery that various of the chemical elements have more than one isotope showed that this assumption was mistaken--that, in this respect, the accommodation demands of the disciplinary matrix were not satisfiable--and that a modification in the use of chemical terminology was required to achieve more nearly adequate accommodation between that terminology and the relevant causal structures.

What is important for our purposes is that the linguistic and classificatory changes which achieved that increase in accommodation--recognizing and naming the different isotopes of the elements, like Carbon 12 and Carbon 14, were not the only response which would have been (exactly just as) adequate to that task. The term "element" could have been reserved for the more fundamental categories we now call isotopes (indeed, this choice might have been closer to the pre-scientific meaning of "element"), and terms could have been introduced for the resulting new "elements," with complementary terminology to represent the relationship which one of these "elements" bears to another if they have the same atomic number. Indeed, the same terms, "Carbon 12," "Carbon 14," "Uranium 235," "Uranium 238," etc. could have been deployed as terms for distinct elements. Accommodation would have been achieved to just the same extent with either modification of the previous linguistic and classificatory practices--either way of denotationally refining the use of the term "element" and the terms for the chemical elements.

We may describe the choice between these two options this way: the programmatic definitions of the term "element" and of the terms for (what chemists then called) "elements" which chemists and others tacitly deployed were mistaken, since the accommodation demands they reflected could not actually be satisfied. In deciding on one or the other of the choices we are discussing, the scientific community would have been tacitly adopting one or another of two different, but equally adequate, ways of appropriately revising those tacit programmatic definitions to bring them into correspondence with the constraints of actual causal structures. Each of these two revisions would lead to different, but equally adequate, conceptions of the explanatory definitions of the relevant terms. Neither way of revising tacit programmatic definitions, and, thus, of revising (explicit) explanatory definitions would have contributed more than the other to accommodation.

The choice which the scientific community actually made was, in that respect, entirely arbitrary. In particular, the preceding bicameral linguistic legislation regarding the use of chemical terminology--which is to say, the previous success at achieving accommodation--did not dictate one or the other of these choices.

Let's simplify the actual history to see this point about linguistic legislation. We can imagine two different counterfactual stories about how reports of (what we call) the discovery of isotopes were written:

Story One: The key paper which reported this discovery began with the sentence, "Many of the 80-odd chemical elements exist in more than one form," and then proceeded to explain the relevant evidence and to introduce the term "isotope." Its publication thereby established the same denotational refinement of the earlier use of chemical terms as the one we have actually adopted.

Story Two: The key paper which reported this discovery began with the sentence, "There are many more chemical elements that the 80-odd previously recognized," and then proceeded to explain the relevant evidence and to introduce terms like "Carbon 12" and Carbon 14" as terms for elements. Its publication thereby established the alternative but equally adequate denotational refinement we have been considering.

Whichever of the two choices had been made, when scientists reported either (a) that there were lots more elements than had previously been recognized, or, if the other choice had been made (b) that the number of elements was 80-odd but not all samples of the same element have the same physical properties because their nuclei may have different numbers of neutrons, they would have been simultaneously reporting new discoveries and establishing new linguistic and classificatory practices appropriate to them. In particular, even though the syntactic structures of the first sentences of the papers in the two different possible stories would suggest that they expressed contradictory propositions, in their respective contexts they would each have been a vehicle for expressing the very same proposition.

The point is quite general. When scientists discover that some of the accommodation demands of their disciplinary matrices cannot be satisfied, and that a change in linguistic and classificatory practice is therefore called for, it is characteristically true that (a) there will be more than one way to effect the required change in linguistic practice such that the choice or one over the others will be arbitrary in the sense just discussed, and (b) when scientists deploy the resources provided by one of these ways of changing existing practices in reporting their discoveries, what they say, or write, is properly understood in that context as both embodying empirical claims and establishing the new linguistic and classificatory practices appropriate to them, so that (c) reports framed in the terms appropriate to different changes in linguistic practice, even though apparently contradictory, would, in the contexts provided by their respective deployments, express the same propositions.

In cases of the sort we have been discussing the choice of terminology in the description of the facts which necessitate denotational refinement <u>constitutes</u> the establishment--in that context--of one rather than another version of the required denotational refinement. Let's call cases of this sort cases of <u>constitutive</u> <u>description</u>.

4.4.1. <u>Constitutive</u> <u>Description</u> <u>of</u> <u>Possible</u> <u>Worlds</u>. I contend that the same sort of constitutive description is a regular feature of our deployment of natural kind terms in describing possible worlds, and that this fact limits the authority of our bicameral linguistic legislation regarding natural kind terms even for world regarding which the identification of the relevant disciplinary matrices for those worlds is unproblematic.

Consider the situation of a philosopher who is exploring counterfactual situations in which there are additional peculiarities of nuclear structure analogous to the actual world peculiarities we have been discussing. She is considering a class of possible worlds in which, in addition to particles with just the same physical properties which actual world protons and neutrons possess, there are also two other sorts of particles which occur in some atomic nuclei. One sort consists of particles which whose physical properties are just like those of actual world protons except that they are very slightly more massive; the other sort consists of particles which are just like actual world neutrons except that they, too, are slightly more massive. She concludes--on the basis of considerations from actual world physics and chemistry (recall that that's what determines the properties of natural kinds in such possible worlds)--that the physical and chemical properties of some substances in the worlds she is considering will be noticeably, but very subtly, different from those of substances in the actual world.

She now sets forth to write up the results of her work. Consider two different ways in which she might choose to describe the class of possible worlds she has been considering:

<u>Choice</u> <u>one</u>: She begins her paper with this sentence: "I here explore the chemistry and physics of possible worlds in which there are two kinds of protons--some of them as massive as actual world protons, and some 1% more massive--and two kinds of neutrons--some of them as massive as actual world neutrons, and some 1% more massive."

<u>Choice</u> <u>two</u>: She begins her paper with this sentences: "I here explore the chemistry and physics of possible worlds in which, in addition to protons and neutrons, there are two other sorts of particles, para-protons, and para-neutrons, 1% more massive that protons and neutrons, respectively."

I contend that her choice is unconstrained by existing linguistic legislation, and that, whichever choice she makes, she will--in that context--be describing <u>exactly</u> <u>the</u> <u>same</u> class of possible worlds. In the relevant disciplinary matrix <u>implemented</u> <u>for</u> one of the possible worlds she is considering, there would be a need for denotational refinement of the usage of the terms "proton" and "neutron". Thus, her description of the features she ascribes to those worlds has just the semantic features of constitutive description in the actual world. Whichever choice she makes, she will be simultaneously describing a class of possible worlds and establishing, in and for the context created by the publication of her paper, the particular form of denotational refinement which she will implement in describing those worlds.

There is no question here of one of her choices being "strictly speaking" more accurate. The programmatic definitions appropriate to the terms "proton" and "neutron" in the actual world are inappropriate for the worlds she is considering, so usage in the relevant disciplinary matrices for those world demands the adoption of one or the other revised programmatic definition for each. Either of her choices will equally well reflect the causal structures of the worlds she is discussing--which is to say that either will allow her to equally well use descriptions of those worlds to convey claims about accommodation and causal structures in the actual world. Therefore, existing bicameral linguistic legislation is silent on the matter. She is free to--indeed she must--legislate by constitutive description.

4.4.2. Constitutive Description and the Limits of Essentialism. Suppose that we now ask, "Is their exact actual world rest mass essential to protons?" and suppose that there is no rest-mass variability among actual world protons. Our philosopher, if she has made choice one in her publishing career, will have already described a possible world in which some protons have a different rest mass, and she could easily publish a sequel in which she described a possible world in which all the protons do. Does this settle the question in the negative?

No, at least not obviously. If she had made the second choice in publishing her paper, our philosopher would have shown that there is an equally good semantics for the term "proton" in non-actual possible worlds which treats proton-like particles with different masses as non-protons. So one could equally well extend the semantics of "proton" in the actual world to other possible worlds in such a way as to make exact (actual world) proton mass essential to protons.

One option would be to say that protons have their actual world mass essentially if and only if there is no way of understanding the semantics of "proton" in different possible worlds (no acceptable constitutive description, for example) which is consistent with the dictates of bicameral legislation and which fails to treat actual world mass as part of the relevant explanatory definition.

This will give a determinate negative answer to the question posed, but it will give rise to a conception of essences for natural kinds quite disconnected from the question of their actual world explanatory definitions. Since it has been the recognition of the importance of a posteriori explanatory definitions which has underwritten much of the return to essentialism in analytic metaphysics in the first place, it is by no means clear what philosophical value there would be in a concern for such essences.

Instead, I propose that the question really doesn't have a determinate answer--given existing linguistic legislation--precisely because either answer can be made (locally) appropriate by a constitutive description which introduces a class of possible worlds in the context of some quite particular scientific or philosophical exposition or discourse.

How can this be? Given that reference to scientifically and metaphysically informative explanatory definitions for natural kinds is essential to an understanding of actual world science, how can it be that there is no correspondingly important metaphysical notion of the essence of a natural kind applicable across all possible worlds? The answer, I suggest, is that the only thing that makes a statement about an (actual world) natural kind in a non-actual possible world true is that it accurately reflects relevant causal facts about the actual world. For any possible world, sufficiently like the actual world that the relevant disciplinary matrix for it to be well defined, but sufficiently different that denotational refinement is required, different token constitutive descriptions which assign different explanatory definitions (different "essences") will each be true in the context established by its use.

What is true is that scientifically and metaphysically interesting essences (that is: explanatory definitions) of actual world natural kinds have their home only in possible worlds very similar to the actual world.

4.4.3. Does Water=H_2O? No. Not, anyway, if the equation is supposed to define water in "all possible worlds." In the first place, the logical form of any explanatory definition (like the definition of water in terms of molecular structure) is not that of an identity statement, but that of a causal explanation. When someone says that water is H_2O, or that water is nothing over and above H_2O, or that "H_2O" is the definition of water, or (if you must) that water is "essentially" H_2O, what she reports is the fact that (a) most of the stuff we call "water" is mainly made of H_2O molecules (b) this fact explain the inductive and explanatory utility of the term "water."

To see that this latter claim does not entail that "Water is H_2O" is true "in all possible worlds," consider again the options open to the philosopher of science described in the last section as she describes the possible worlds in which there are two sorts of proton-like particles and two sorts of neutron-like particles. Suppose that the adopts choice two and (constitutively) describes those worlds as containing "para-protons" and "para-neutrons."

Imagine now that she wishes to consider the sub-class of the possible worlds in question in which the only nomologically possible nuclear structures are those which either (a) contain just protons and neutrons,, or (b) contain just para-neutrons and para-protons. She will face another opportunity for constitutive description: she may (as she wishes) describe these worlds as either (a) worlds in which the chemical elements come in more different forms than they do in the actual world, since every actual world isotope of an element is matched by an additional form in which the nuclei are composed of para-protons and para-neutrons, or (b) worlds in which there are 208 elements instead of the actual world's 104: hydrogen and para-hydrogen, carbon and para-carbon, etc.

Suppose that she makes the latter choice and that she wishes to report the fact about actual world water, that the exact mass of protons does not matter very much to the biochemistry of water. Once again, she faces as opportunity for constitutive description. She could report this fact by saying either (a) that in the worlds she is considering the "two kinds of water" (the mono-oxide of hydrogen and the mono-oxide of parahydrogen) have almost identical biochemical properties, or (b) that in those worlds "water" (the mono-oxide of hydrogen) and "para-water" (the mono-oxide of para-hydrogen) have almost identical biochemical properties.

The point here is that, in the relevant disciplinary matrices for such worlds, the term "water" would require denotational refinement (since hydrogen mono-oxide and para-hydrogen mono-oxide will differ in some physically important ways) and neither way of achieving that refinement would be more or less in accordance with our bicameral linguistic legislation than the other. No metaphysical, or physical, or chemical or linguistic facts dictate that we say that water is H_2O in all possible worlds even though the formula "H_2O" picks out the real essence of water in the actual world.

4.5. Metaphilosophical Consequences. In contemporary analytic philosophy we routinely consult our modal intuitions about natural kinds, epistemic categories, moral properties, secondary qualities, aesthetic attributes, and the like in "all possible worlds," and we take the results of this sort of consultation to be metaphysically informative. As it happens, our modal intuitions often lead us to have quite definite and quite strong convictions about such matters in possible worlds very much unlike the actual world.

If the points made in the present part of this paper are right, these convictions are almost always wrong. Our (strong) modal intuitions by far outrun the scope of the bicameral legislation which establishes truth conditions for the modal statements we consider. They represent the methodologically inappropriate intrusion of a prioristic elements into areas of a posteriori science and a posteriori metaphysics.

I do not mean to deny that intuitions (philosophical and scientific) have a legitimate methodological role. They are often reflections of important tacit knowledge (Boyd 1988). But our modal intuition, I suggest, are reliable only about very nearby possible worlds.

5. The Politics of Scientific Knowledge.

5.0. Ideology and Anti-Realism. Among the many and diverse motives for the relativist and anti-realist approaches to scientific knowledge which characterize much contemporary work in the humanities, none is more admirable than a concern to come to grips with the role of science, and of conceptions of scientific expertise and objectivity, in social ideology--especially in rationalizing oppressive or unjust social and political arrangements. It is a sad fact that, at least from the middle of the last century on, scientific findings and the authority of scientific experts have been available for the rationalization of oppressive policies and social structures. This phenomenon is, perhaps, clearest in human biology where there has been a systematic tendency for the findings of respectable experts at any given time--in genetics, say, or the psychological theory of individual and group differences--to ratify whatever relations of disproportionate power and wealth prevail at that time. Racist justifications of colonial policy or chattel slavery in the 19th century or of social, economic and political inequalities or oppression in the 20th are paradigm cases.

This phenomenon makes the very notions of scientific objectivity and expertise into weapons of oppression, and many "postmodern" critics of science have intended their relativist and anti-realist conception as contributions to a critique of the ideological role of science.

Those of us who think that the sciences sometimes deliver real knowledge of real phenomena (and sometimes discover the essences of things) need to have something informative to say about such matters. In this case, as in so many others, it turns out that we have something to learn from the slogans of the other side. There is, for example, a perfectly good sense in which the epistemology of science is a (partly) political matter.

5.1. Institutions, Expertise and Theory Testing. We have already seen that, to a good first approximation, scientific methodology dictates that a theory receives significant evidential support from a body of successful predictions (or apparently successful explanations) just in case those predictive or explanatory successes favor the theory over its projectible rivals, where projectibility is a matter of plausibility in the light of established background theories. We have seen, furthermore, that the applications of this methodological principle in practice are epistemically reliable only if (a) the propositional content of the relevant background theories is close enough to the truth that an answer near the truth is typically among the projectible answers to any given scientific question and (b) the social and political structures of scientific institutions and the surrounding society are such that these projectible answers near the truth tend to get articulated, developed, and publicized in sufficient detail to make their merits clear.

Consider how the methodology we are considering is implemented in practice. No scientist or research group has the expertise by itself to evaluate the vast majority of the background scientific theories upon which their judgments of projectibilty depend. Instead, they have no choice but to rely on (other) experts in the relevant discipline. Indeed, no individual scientist or research group has the expertise (or the time) to undertake an evaluation of the expertise of those other experts to whose expertise they are thereby deferring. Instead, they must to a very great extent rely on scientific institutions (graduate programs, prestigious research institutions, prestigious journals, academies of science, etc.) to certify experts and expert opinions.

Thus, again to a very good first approximation, the scientific methodology for assessing a proposed theory is to ask (a) whether it is theoretically plausible given the background theoretical assumptions underwritten by the opinions of experts licensed as experts by scientific institutions, and (b) if it is, whether or not is proves superior, in the light to observational or experimental testing procedures similarly underwritten by licensed expertise, to the competing theories which are similarly theoretically plausible. Individual scientists or research groups may, of course, challenge the credentials of some of the licensed experts, or they may contest some of the findings of those experts, but even these critical activities must take place in a methodological context largely determined by the relevant scientific institutions. It could not be otherwise, given the complexity of the world, the cognitive limitations of individual researchers and research groups, and the need to achieve theory-dependent accommodation.

There is no other method of science; there is no epistemologically revealing idealization which makes the reliance on institutions go away. If we are naturalists about the epistemology of science, then we must conclude that scientific knowledge--when we have it--is a matter of true (or approximately true) beliefs reliably produced by these institution-as-well-as-theory-dependent methods. So, the epistemic reliability of scientific methods--when they are reliable--is a matter of the tendency of scientific institutions to certify as projectible, and to publicize, some approximately true answers to the questions scientists address.

5.2. Institution and Ideology. We are now in a position to pose the problem about ideology and scientific expertise which concerns many admirably motivated postmodern thinkers. If you look at the history of ideology in biology--the history of racism in biology for example--a striking fact emerges (see Gould 1981 for a good discussion). The very same scientific institutions (prominent university departments, prominent journals, academies of science, and all that)--and indeed the very same prominent scientists--which (who) reliably tended to certify as projectible some approximately true answers to many sorts of questions in human biology and non-human genetics, have reliably and systematically tended to treat as non-projectible--indeed as unrespectable--anti-racist (and true) answers to questions about human individual and group differences.

If the methods of science were theory-and-institution independent, then this phenomenon could be diagnosed as a matter of the corruption of scientific methodology by the intrusion of political power and culture. But--as many postmodern thinkers have recognized--this response in unavailable. The very institutions which define scientific methodology are the ones whose licensing of experts and projectibility judgments has usually sustained (what we later recognize as) racism in science.

That's why the notions of scientific expertise and objectivity are such powerful political weapons for oppression, and their political effects are magnified because most of the scientists involved in the production of (what is, in fact) racist science are, and are known to be, conscientiously practicing scientific methodology and have, and are known to have, no particular racist intentions or agendas. From the methodological perspective available to them, what seems in retrospect transparently ideological science looks no different from those aspects of their scientific work which we, in retrospect, do treat as genuine achievements.

5.3. Realist Reliablism and the Limitations of Scientific Method. No realist naturalism can afford to ignore these cases in which scientific methodology "error-tracks" instead of "truth-tracks". Philosophers of science--and epistemologists generally--look to the methods of science as crucial examples of epistemically reliable practices. This is a good idea, but only if one's understanding of the reliability of those methods is properly informed by an epistemological naturalism which is itself informed by an accurate appraisal of the history of science. Even in mature sciences in which much of the propositional content of relevant background theories is approximately accurate and insightful, scientific methods are not always generally reliable. Their reliability depends as well on facts about politics and power within both scientific institutions and the relevant political and social order.

In particular, there is--in many societies--a systematic tendency for scientific "findings," in the human sciences, genetics, animal behavior, etc. to be such as to ratify existing patterns of power and social stratification. Given the structure of those societies, the relevant "finding" are almost always systematically and deeply false, because they portray temporary patters in human society as biologically inevitable.

It follows that--even in mature sciences--the reliability of methodology depends--in ways that have political as well as technical explanations--on what questions are being addressed, and the pattern of this dependence cannot be adequately understood except in the context of a political understanding of the structure of scientific institutions and of their societal setting. Since no epistemically informative naturalistic account of scientific knowledge can be formulated which ignores the role of scientific institutions, the epistemology of science is necessarily a (partly) political discipline. Naturalistic epistemologists of science who ignore the political dimensions of scientific institutions tacitly subscribe to historically and politically naive conceptions of those institutions, and their epistemology suffers in consequence.

5.4. Error Correction and the Epistemology of Science. I want to respond to an obvious objection to what I have just said. We know about the ideological functions which appeals to science and to scientific expertise can play because we now know, about some of the ideologically determined "findings" of earlier scientists, that they were mistaken and that the mistakes involved were just such as to ratify existing patterns of power and stratification. But, the objection goes, we only know that these earlier conception were in error because later scientific studies showed that this was so. So scientific methodology tends, in the long run, to self-correct, even when the relevant errors are ideological. The impression that the methods of the mature science are not always generally reliable stems from looking at the issue of reliability on too short a time scale.

I want to respond using an example of "self-correction" about which I have some personal knowledge. Gould (1981) describes an institutionally situated research tradition in genetics, anthropology and the psychology of human differences which has persisted since the middle of the last century and which has always centrally involved the articulation of racist and class prejudiced theories of individual and group differences--theories which, in fact, served to justify stratification, colonial policies and wars. The publication of Gould's book was one of the later stages in a (sadly temporary) period of sustained and persuasive scientific critique of just these racist theories (see Block and Dworkin, 1976 for a representative sample of some earlier work in this same critique). The appearance of this sustained scientific critique of racist theories of intelligence is, presumably, an example of the sort of self-correction of ideological error in science to which the objection refers.

The scientific and philosophical work that went into the critique of racist theories of intelligence during this period involved a number of important theoretical, methodological, statistical and philosophical re-evaluations of the research which underwrote those studies. There were also a number of important empirical studies (largely correlational studies) of the relation between various alleged measures of intelligence and socio-economic and cultural variables. Of course all of these various studies deployed all the latest methodological, philosophical, genetic and statistical resources.

It remains true, however, that the scientific and philosophical critiques did not depend in any fundamental way on new innovations in genetics, in psychology, in statistics, in sociology, or in philosophy. Approximately the same critiques could have been launched using the scientific, philosophical and statistical resources available to scientists and philosophers at any time since the mid-1930's. But, they were not.

What had changed by the late 1960's and the 1970's when the scientific critique of this sort of scientific racism was undertaken was, instead, something about the imaginative capacities of many scientists and philosophers: their capacities to imagine and articulate alternatives to the racist explanations for patterns of work and behavior, and for relations of power and wealth. Some significant number of scholars deployed the resources of these alternative explanations to frame criticism of both the methods and the findings of earlier racist studies of individual and group differences. It is important that even these alternative explanations were not new, and their articulation relied very little on new theoretical or conceptual resources. What changed was the number of researchers with scholarly status who found them sufficiently projectible to take seriously.

We may ask where the work was done which resulted in the new insights that underwrote these improved projectibility judgments. It takes no credit away from the scholars who contributed to the critique (often at considerable risk to their academic careers) if we acknowledge that most of it was not done in Cambridge, or in Berkeley, or in Ann Arbor. It was done instead in Delhi and Calcutta, in Dien Bien Phu, Saigon and Hue, in Johannesburg and Pretoria, in Watts and Harlem. it was done, largely but not exclusively, by militant people of color; most often by disciplined, organized people of color with guns. Their struggles against racism and imperialism made salient important non-racist alternatives to racist explanations of power and stratification. The intelligence, strategic and tactical skills, moral insights, and courage they often enough displayed in the course of these struggles provided challenges to the propositional content of racist theories.

All of these factors produced the changes in (some scholars') projectibility judgments, and enabled the readers of their critiques to more readily appreciate their cogency, but they were not factors which were, in any obvious sense, part of scientific enterprises or institutions. In fact, those institutions have proven remarkably resistant to incorporation of the resulting insights. The critique of racism in theories of individual and group differences did not produce the sort of long-standing changes in institutionally sanctioned research methods which many expected, with the result, in the United States at least, that there is now a significant academic industry in which scientists publish what amount to recycled racist (and sexist and class prejudiced) arguments from the 1960's.

In this important case, science did not self-correct. Some scientists (and philosophers, as it happened) made a significant, but still minor, contribution to the (temporary, as it happens) modification of scientific thinking on matters of individual and group differences in a less racist, less ideologically driven direction. They should be praised, but the scientific method by itself was not adequate to the epistemological task of rectifying (even temporarily) the errors in scientific practice. That task was an essentially political one, which is not to say that it was not at the same time an epistemological task. Epistemology is partly a matter of political understanding and (if epistemology is to be sometimes an applied discipline) of political action. Any adequate realist naturalism about the epistemology of science must acknowledge this fact.

5.5. Rethinking Science. One of the expectations of philosophers (and others) who look to the sciences as a source of epistemic insight is that the methods of the mature sciences will be, often enough, generally reliable. We have just seen that this expectation--this accommodation demand of scientific epistemology--is met only in a certain qualified way by the actual epistemically relevant causal features of what we ordinarily recognize as scientific practice. When this sort of thing happens, one response within a disciplinary matrix is often to revise standards of classification in such a way that the accommodation demands are more fully satisfied.

In cases like that of racist ideology in science, the standard approach along these lines is to deny that racist science is really science. The results of the previous sections indicate why this approach is unpromising: the institutional and theoretical structures which govern epistemically reliable science are too closely bound up with those which licensed (what we now recognize as) racist science.

If it is unworkable to narrow the (explanatory) definition of science to exclude ideology in science, the option remains of making the definition more inclusive so as to include the practices which did, in fact, contribute to the (perhaps temporary) refutation of scientific racism. Perhaps we should count systematic political struggles against racism and imperialism as part of scientific activity. If, say, Vietnamese peasants in their anti-imperialist struggles contributed to overturning scientific racism, perhaps we should count those struggles as part of the history of science.

I'm not at all sure that this proposal is ultimately appropriate. What I am sure of is that there is much more to be said in its favor than one might have thought. In the first place, it is important to recognize that, in their systematic political--and military--struggles against various forms of racist and colonial oppression the militant activists of the post-war period necessarily sought, using methods recognizably like those of the empirical sciences, to understand the economic, historical and political roots of their oppression in order to overcome them. Counting this understanding as scientific even though it was motivated by concerns for application is, in that regard, on a par with recognizing--as we should--that lots of real science gets done by people we classify as engineers rather than as scientists.

Once we have seen this, we can identify other reclassifications with a similar flavor: treating the research activities of "amateur scientists" (whether 18th and 19th century naturalists or 20th century amateur astronomers) as part of science. In these cases, as in the cases of engineers and Third World peasants, we can discern one reason why the proposed reclassifications seem counterintuitive: they erase distinctions of status associated with the terms "science" and "scientist." This is not obviously a bad thing.

One additional consideration is crucial here. Not only did the anti-racist and anti-imperialist struggles we are considering contribute indirectly to the scientific critique of scientific racism, they typically reflected, on the part of their participants, a more accurate understanding by far of the political economy of racism and colonialism than that provided by institutional social sciences. If we should choose to treat the Vietnamese peasants who fought against the French and the U.S. as having done social science, then they surely did better social science than most mainstream social scientists.

It may still be true that the inductive/explanatory roles played by our uses of the term "science" would make the proposed reclassification inappropriate. Whatever may be the case in that regard, the fact remains that, in roughly the period from 1945 to 1985, the scientifically, politically and morally most important insights about the human sciences came from the work of marginalized people of color and their allies in struggles against racism and imperialism, and from related efforts of other marginalized people inspired by those struggles. This is the sort of fact celebrated by some postmodernist "standpoint theorists" in epistemology. No adequate realist naturalism about scientific knowledge can afford to ignore it.

Richard Boyd
Cornell University

Bibliography

Block, Ned J, and Dworkin, Gerald, eds. 1976. The IQ Controversy. New York: Pantheon.

Boyd, R. 1982. "Scientific Realism and Naturalistic Epistemology." in P.D. Asquith and R.N. Giere (eds.) PSA 1980. Volume Two. E. Lansing: Philosophy of Science Association.

______. 1983. "On the Current Status of the Issue of Scientific Realism." Erkenntnis 19: 45-90.

______. 1985a. Lex Orendi est Lex Credendi. in Churchland and 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: MIT Press.

______. 1988. "How to be a Moral Realist." in G. Sayre McCord (ed) Moral Realism. Ithaca: Cornell University Press.

______. 1989. "What Realism Implies and What It Does Not" Dialectica.

______. 1990. "Realism, Conventionality, and 'Realism About'" in Boolos, ed. Meaning and Method. Cambridge: Cambridge University Press.

______. 1991. "Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds." Philosophical Studies 61: 127-148.

______. 1992. "Constructivism, Realism, and Philosophical Method." in John Earman, ed. Inference , Explanation and Other Philosophical Frustrations.

______. 1993. "Metaphor and Theory Change" (second version) in A. Ortony (ed.) Metaphor and Thought, 2nd Edition. New York: Cambridge University Press.

______. forthcoming a. "Kinds, Complexity and Multiple Realization: Comments on Millikan's 'Histori cal Kinds and the Special Scienc es'" to appear in Philosophical Studies.

______. forthcoming b. "Homeostasis, Species, and Higher Taxa," in R. Wilson, ed. Species: New Interdisciplinary Essays. Cambridge: MIT Press.

Field, H. 1973. "Theory Change and the Indeterminacy of Reference." Journal of Philosophy (70): 462-481.

Ghiselin, M. 1974. "A Radical Solution to the Species Problem," Systematic Zoology (23): 536-544.

Goodman, N. 1973. Fact Fiction and Forecast, 3rd edition. Indianapolis and New York: Bobbs-Merrill.

Gould, Stephen Jay 1981. The Mismeasure of Man. New York: Norton.

Hull, D. 1965. "The Effect of Essentialism on Taxonomy--Two Thousand Years of Stasis," British Journal for the Philosophy of Science 15, 314-362 (part one) and 16, 1-18 (part two).

_______. 1978. "A Matter of Individuality," Philosophy of Science (45): 335-360.

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

Kripke, S.A. 1971. "Identity and Necessity." in M.K. Munitz (ed.) Identity and Individuation. New York: New York University Press.

__________. 1972. ""Naming and Necessity." in D. Davidson and G. Harman (eds.) The Semantics of Natural Language. Dordrecht: D. Reidel.

Mayr, E. 1988. Towards a New Philosophy of Biology Cambridge: Harvard University Press.

Putnam, H. 1972. "Explanation and Reference." in G. Pearce and P. Maynard, eds. Conceptual Change. Dordrecht: Reidel.

_________. 1975a. "The Meaning of "Meaning'." in H. Putnam, Mind, Language and

Reality. Cambridge: Cambridge University Press.

________. 1975b. "Language and Reality." in H. Putnam, Mind, Language and Reality. Cambridge: Cambridge University Press.

__________. 1978. "Realism and Reason," in Putnam, Meaning and The Moral Sciences. London and New York: Routledge and Kegan Paul.

__________. 1980. "Models and Reality" Journal of Symbolic Logic 45, 464-482.

_________. 1983. "Why There Isn't a Ready-Made World." in H. Putnam, Realism and Reason. Cambridge: Cambridge University Press.

Quine, W. V. O. 1969. ""Natural Kinds." in W.V.0. Quine, Ontological Relativity and Other Essays. New York: Columbia University Press.