Scientific Realism and Naturalistic Epistemology[1]

Richard Boyd

Cornell University

1. Introduction

By "scientific realism" philosophers ordinarily mean the doctrine that non-observational terms in scientific theories should typically be interpreted as putative referring expressions, and that when the semantics of theories is understood that way ("realistically"), scientific theories embody the sorts of propositions whose (approximate) truth can be confirmed by the ordinary experimental methods which scientists employ. There are as many possible versions of scientific realism as there are possible accounts of how "theoretical terms" refer and of how the actual methods of science function to produce knowledge.

What I will do in this paper is to explore the consequences of one such version of scientific realism, a version which embodies the implicatures as well as the implications of the realist slogan that reality is prior to thought. What I have in mind is a dialectical and naturalistic conception of how scientific language works and how scientific knowledge is achieved, a conception according to which not only scientific knowledge, but the language and methods of the sciences as well, represent hard-won victories in a continuing struggle to accommodate our intellectual practices to the structure of an independently existing world. In broad outline, the picture of science which this conception presents goes like this:

In the first place, the world has a quite complicated causal structure, and many of its most important features are unobservable to the unaided senses.

Scientific knowledge extends to both the observable and the unobservable features of the world, but it is achieved by a process of successive approximation: typically, and over time, the operation of the scientific method results in the adoption of theories which provide increasingly accurate accounts of the causal structure of the

PSA 1980, Volume 2, pp. 613-662

world. If we think of beliefs or theories as being "accommodated" to the world insofar as they are accurate descriptions of some of its features, then scientific knowledge procedes by accommodation by successive approximation.

What is true of scientific knowledge is also true of scientific language as well. Adequate scientific terminology must provide us with the descriptive machinery necessary to describe the (typically unobservable) fundamentally important features of natural phenomena, and to classify them in ways which reflect the complex causal properties which these phenomena possess. Scientific language must provide us with the descriptive machinery necessary to "cut the world at its joints". This sort of accommodation between scientific terminology and the causal structure of the world is, like scientific knowledge, achieved by successive approximation. Moreover, this accommodation is not merely a matter of the introduction of new terms to reflect new discoveries and the deletion of terms which reflect the influence of subsequently refuted theories. Nor is the process of accommodation merely a matter of these processes together with the progressive refinement of usage of existing terminology (though that too is important). The very mechanisms of reference--the ways in which scientific terminology is connected to features of the world--undergo a development--typically in the direction of a closer and "tighter" fit between scientific terminology (in use) and the important causal features of reality. Reality is prior to thought not only in that its structure is largely independent of what we believe, but also in that the very machinery of thought (or, at any rate, of the public expression of thought) undergoes continuous accommodation to the structure of a largely independent causal reality.

Not only do theories and language accommodate to the world by successive approximation: so do the scientific methods and epistemological principles by which knowledge is achieved. Methods in particular sciences are theory-dependent, and they become more closely accommodated to the structure of the world as the theories upon which they depend become more accurate. Moreover, not only the methods of particular sciences but also the more general features of scientific or experimental method develop by successive accommodation to the causal structure of the world. No inductive method possesses *a priori* justification. For any significant general feature of the scientific method, it is possible to imagine possible worlds in which it would be inappropriate, but in which some alternative methodological strategy would make the acquisition of knowledge possible. The methods characteristic of scientific rationality in the actual world reflect the progressive accommodation of our methodological practices and epistemological standards to the particular structure which our world possesses. Reality is prior to thought not only with respect to the correctness of theories and the appropriateness of the language in which they are expressed, but also with respect to the standards by which the rationality of thought is to be assessed.

Finally, the accommodation of theory, language and method to the world is not only a matter of successive approximation, it is also a profoundly dialectical process. In the first place, the general tendency of accepted scientific theories over time to become progessively more accurate depends on the fact that later scientific theories are, almost always, refinements or modifications of earlier theories in the light of new evidence or new theoretical considerations. The effect of such modifications is (again, typically and over time) to preserve and extend the grains of truth in preceding theories while eliminating errors. Similarly, the process of accommodation between scientific language and the structure of the world reflects a dialectical process of refinement and modification of linguistic practice, and a corresponding dialectic operates in the development of scientific methodology and our understanding of it. Moreover, these three processes of accommodation display a dialectical relationship of mutual dependence: new theoretical knowledge leads to improvements in scientific language and in methodology; better methodology leads to greater theoretical knowledge, and so forth.

If this picture of science is approximately accurate, then an adequate philosophy of science must be realistic since it must reflect the fact that knowledge of "theoretical entities" is possible. It must also be naturalistic (in the sense that it sees not only theoretical issues, but linguistic and epistemological issues as well, as broadly empirical issues depending for their solution on _a posteriori_ considerations); finally, it must be dialectical, in the sense just discussed. Now, the picture of science which I have just sketched is --at least in its pre-analytic formulation--commonsensical. (I do not mean to say that it is _the_ commonsense view of science; what is commonsensical is a matter of the current climate of opinion and there is no doubt that some skeptical mixture of logical positivism and relativism is also now commonsensical.) Nevertheless, this version of scientific realism--if it is developed and articulated into a systematic position in the philosophy of science--has, I believe, some important but controversial implications for our understanding of scientific knowledge.

In this paper I will explore some of these implications which bear on outstanding issues in the philosophy of science. I will also say something about what _this version_ of realism does not imply; I am persuaded that the attractiveness of anti-realistic positions in the philosophy of science often stems as much from the belief that realism has unacceptable consequences as from the (very powerful) verificationist epistemological considerations which are the traditional basis for the rejection of realism. In general, I will content myself with rather brief discussions of the doctrines I shall defend. I have defended most of the doctrines elsewhere, and, often, so have a number of other philosophers. The reader who wishes to see a fuller defense of these doctrines may consult the references cited.

2. Outline of a Naturalistic and Dialectical Version of Scientific Realism

2.1. Nagel's Dictum, The Theory Ladenness of Method and the Naturalistic Epistemology of Science

One dimension along which we may discern differences between versions of scientific realism concerns the extent to which their general accounts of scientific methodology is compatible with that of nonrealists in the empiricist tradition. One view of the matter is that realists and, say, instrumentalists share a common basic conception of the logic and methods of science, but that they disagree about whether those methods are adequate to establish knowledge of unobservable phenomena. This conception is stated with admirable clarity by Nagel (who also concludes--as a realist of course would not--that the disagreement between realists and logical empiricists about the semantics of scientific language and the extent of scientific knowledge is merely verbal). "It is difficult to escape the conclusion that when the two opposing views on the cognitive status of theories are stated with some circumspection, each can assimilate into its formulation not only the facts concerning the primary subject matter explored by experimental inquiry but also the relevant facts concerning the logic and procedures of science. In brief, the opposition between these views is a conflict over preferred mode of speech." (Nagel 1961, pp. 151-152).

An alternative realistic position might hold that realists and their empiricist opponents agree about the methodology by which instrumental knowledge is obtained, and that they agree in believing that _this_ methodology is insufficient to establish theoretical knowledge. Realists, on this view, must propose additional epistemological or methodological principles to justify their claim that theoretical knowledge is possible. This position _seems_ to be the one which J.J.C. Smart adopts; he appears to think that the additional principles are philosophical principles rather than principles of scientific evidence (Smart 1963, Chaps. I and II). Of course, one might hold that the additional methodological principles were also, in whatever the relevant sense, principles of experimental methodology. If I am right, none of these approaches to the epistemology of science is correct.

Let me introduce some terminology: By the "instrumental reliability" of a scientific theory I will mean its ability to provide (given suitable "auxiliary hypotheses") approximately accurate predictions about the behavior of observable phenomena. By "instrumental knowledge" I will mean the knowledge about particular theories that they are instrumentally reliable, and the concomitant knowledge about observable phenomena. By the "instrumental reliability" of methodological principles, I mean their capacity to contribute to the production of instrumental knowledge.

Now it is almost uncontroversial among philosophers of science that instrumental knowledge is not only possible but actual. Similarly, it is almost uncontroversial that (at least some of) the actual methods of science are instrumentally reliable. Moreover, these claims are not merely matters of _current_ consensus. They have been presuppositions of the philosophy of science (and, importantly, of empiricist philosophy of science) from the beginnings of the discipline. One of the interesting effects of these presuppositions has been to introduce into the philosophical works of logical positivists and their successors an empirical or naturalistic strain entirely out of keeping with their oft proclaimed allegiance to the distinction between empirical inquiry and the sort of logical and conceptual inquiry properly characteristic of philosophy. By "rational reconstruction" positivists meant the task of identifying and explicating the sound features of actual scientific methods, and the well-confirmed features of actual scientific theories. Rational reconstruction, in the hands of logical positivists, has two distinct components: careful examination of the actual methods and findings in the sciences, and the application to those methods and findings of some version or other of anti-metaphysical verificationist principles. While the second of these components was supposed to represent purely conceptual and logical considerations, the first was evidently tied to detailed consideration of the work of actual scientists.

Furthermore, it was by no means the case that the first of these components was always subsidiary to the second (as it might have been if the examination of actual scientific practice had been a mere "heuristic" device to aid in the discovery of methodological or conceptual principles which would later be defended on purely conceptual and logical grounds). Instead, it was (and is) utterly routine for the results of *aprioristic* reasoning concerning the foundations of science to be abandoned when found to be incompatible with actual scientific practice (for an appeal to such considerations see Hempel 1965, chapter 4; in fact, considerations of this sort have been the decisive force in the development of recent philosophy of science). Indeed, the intrusion of such precursors of naturalistic epistemology into empiricist philosophy of science was not limited to features of philosophical method. Causal theories of measurement analogous to causal theories of perception (and, to some extent, causal theories of reference) were recurring, if obscure, themes throughout the 1950's, especially in the work of Feigl (see Feigl 1956; see also MacCorquodale and Meehl 1948, which was obviously strongly influenced by Feigl, and in which the doctrines of this section are anticipated).

I have elsewhere defended (see Boyd 1972, 1973, 1979, 1980, forthcoming (a), forthcoming (b)), the view that the naturalistic elements of the strategy of rational reconstruction actually provide the basis not only for the defense of scientific realism against various versions of logical empiricism, but also for the articulation of an account of the epistemology of science which represents a fundamental break with the empiricist tradition, and with the conception of realists' epistemology represented by Nagel's dictum and by the alternative offered by Smart. What I have sought to establish is the following claim: No scientifically plausible explanation of the *instrumental* reliability of actual scientific methods is possible which does not portray those methods as reliable for the acquisition of theoretical knowledge as well. Moreover, the reliability (instrumental or theoretical) of scientific methods at a given time will typically be explicable only on the assumption that the existing theoretical beliefs which form the background for its operation are (in relevant respects) approximately true. The basic idea which I have defended is that theoretical considerations are so heavily and so crucially involved in the operation of actual scientific method that the only way to explain even the instrumental reliability of that method is to portray it as reliable with respect to theoretical knowledge as well.

In order to indicate how this view can be defended I want to examine the epistemological role of three of the many sorts of ways in which theoretical considerations influence our scientific practice. It will be useful in this regard to introduce two additional technical terms.

I will call two theories T_1 and T_2 observationally equivalent with respect to an existing body of accepted scientific theories (i.e., with respect to an existing "total science") if the same observational consequences would follow from each of the two following "total sciences":

(a) the existing total science modified as it would be by the adoption of T_1

(b) the existing total science modified as it would be if T_2 were adopted.

Given any plausible initial total science there will be infinitely many equivalence classes under this relation. One way of putting the problem of the instrumental reliability of scientific method is this: In fact we choose one of these equivalence classes each time we accept a theory, and we do so on the basis of finitely many observations. So some other criteria than consistency with observational data must be at work. Call these the "extra-experimental" criteria. Whatever these extra-experimental criteria are, they work. In the long (but not very long) run we get quite good predictive theories. Why do these criteria work?

Let me first discuss the most commonplace sort of extra-theoretical criterion which philosophers of science have recognized [roughly, entrenchment]: We, in fact, take seriously only those theories which relatively closely resemble our existing theories in respect of their ontological commitments and the laws they contain. We prefer theories which quantify over familiar "theoretical entities"--or at least entities very much like familiar ones (or, in some cases, appropriate constituents of familiar entities); we prefer theories which predicate of theoretical entities familiar properties--or at least properties like familiar ones; we prefer new theories whose laws are--if not consistent with those we have previously adopted--at least compatible with the maintenance of most of our previously accepted laws as approximations. Generally, we reject outright any proposed theory which contradicts the laws we consider best confirmed unless a real crisis is at hand--and even then we will strongly prefer new theories which preserve the old laws as approximations.

Two points about these extra-experimental criteria are important. In the first place, these criteria of preference for theoretically plausible theories are the real basis for the judgments which are traditionally glossed as judgments of "simplicity": our preference for "simpler" theories is, in the first instance, a preference for theories which represent relatively "simple" modifications of our existing theories. Moreover, this notion of simplicity has an epistemological component: we prefer to preserve those features of existing theories which seem best confirmed and to accept changes in features whose evidential status is less secure. In particular, "simplicity" is not a theory neutral notion; simplicity judgments are profoundly dependent on the existing theoretical tradition, and they rest upon epistemological judgments about the "degree of confirmation" of various components of that tradition, judgments which are--as we shall see--themselves theory-dependent.

A second point, which has been emphasized by Putnam, is the following: In the course of scientific research about some particular issue, only a small handful of theories are "in the field" at any one time. If this weren't so, our research efforts would not be so narrowly and carefully focused as they are. "Simplicity" judgments (that is, judgments of theoretical plausibility) form the criteria by which the field is narrowed. One consequence of this narrowing, is that only a few of the infinitely many equivalence classes of alternative modifications to our existing total science are taken seriously with respect to any given issue. It is from this small handful that, by judicious observation or experiment (where judiciousness is also a theoretical notion), we make our choice. We ignore infinitely many (equivalence classes of) alternative modifications of our existing total science, which have never been tested (much less refuted) by direct experiment. And, we get away with it! This "narrowing down" of our options seems to contribute to the instrumental reliability of the scientific method, rather than to detract from it, as one might expect. Why? Why is this strategy reliable with respect to the task of finding instrumentally reliable theories?

Another theory-laden methodological principle which is almost as commonplace is that which countenances theoretical criticism, modification or extension of procedures of measurement and detection for "theoretical" entities and magnitudes. This principle has, if anything, even more striking instrumental consequences.

Suppose that T_1 (t) is a well confirmed theory containing the theoretical term t, and supported by observations in some class D. Suppose that "measurements" of t have thus far been possible only using measurement procedures $m_1, \ldots, m_r$ whose reliability is asserted by "mini-theories" $M_1, \ldots, M_r$.

Now imagine that some new theory $T_2(t)$, only distantly related to T_1, is confirmed by entirely different observations. T_2 has the happy consequence, M_{r+1}, that a new procedure m_{r+1} is suitable for the "measurement" of t under circumstances well outside the range of application of $m_1, \ldots, m_r$.

Under these circumstances, if T_1 and T_2 are each quite well supported by experimental evidence, we shall confidently expect T_1 (t) to yield true (or approximately true) predictions when employed with the new measurement procedure m_{r+1}.

And we get away with it! Why? What is "measurement", and what is its relation to theory, which would permit us to be confident that $(T_1 \wedge M_{r+1})$ yields true observational predictions _even if no non-trivial observational prediction from this conjunct has ever been tested before_? [Actually, as I will suggest later, this puzzle about measurement and its theory-dependence is a special case of a general puzzle about "unity of science" and the epistemological role of univocality judgments regarding theoretical terms.]

Finally, consider the question of experimental design: Suppose that T is a suitably plausible theory. Which experimental tests are sufficient to warrant our accepting it, and expecting its observational predictions to be approximately true? Which finite (and typically small) number of experimental tests can we count as suitably representative for an assessment of the predictive reliability of T? [Remember, just to make it more interesting, that we are actually assessing the reliability of T taken jointly with other--_perhaps not yet discovered_--well-confirmed auxiliary hypotheses.] As an answer, I propose the following:

The Fundamental Rule of Experimental Design:

1. Subject T to _theoretical_ criticism. [Ask, in the light of the best available theories, what alternatives there are to the mechanisms/processes posited or required by T. What mechanisms, known on the basis of other theories, might interfere with the operation of the mechanisms which T posits or requires? Does the plausibility of T depend upon theoretical considerations which themselves rest on a theory now in dispute? What weakness might that indicate in T? Etc.]

2. After you have subjected T to theoretical criticisms [this is, of course, typically a public rather than an individual activity], then test T under circumstances representative of those _which theoretical criticism indicate as places where it might plausibly go wrong_.

That is how we do it. And that's as theory-dependent as you can get. And it works--the very success of our practice indicates that the use of theoretical criticisms enables us to pick finitely many experiments which are sufficiently representative, so that the theories which we do accept turn out to be very good predictive instruments indeed.

Why does _this_ principle work?

Before we answer these three questions, there is a neat point to be made. The first and the third of these principles are related in that both require that a theory be tested against plausible rivals. Some philosophers have treated the first principle as purely pragmatic. We are only able to think up a few theories, which, in turn, seem natural to us because they resemble theories we have already got. We test these out first, not because we have reason to believe that one of them will work, but because it is pragmatically sound to test those theories you already have first. It is like picking out a hammer--you see if the ones at the local hardware store work before ordering something more esoteric.

We can already see that this pragmatic justification leaves a central issue unsettled: since we almost never send away for one of the esoteric theoretical hammers (of which there is an infinite variety), why do we so often hit the nail on the head?

Better yet, however, is the observation that the third principle requires that a proposed predictive instrument be tested against plausible rivals, even when those rivals are mere "hunches". We may have an alternative to a proposed theory T which suggests that it might go wrong under experimental circumstances C, even though our alternative makes no _specific_ prediction about C whatsoever. The relevant alternatives to T need not be predictive instruments _at all_! Theory testing is very much unlike hammer buying.

A satisfactory naturalistic answer regarding the instrumental reliability of each of these methodological principles is available _if_ one assumes that they apply in a situation in which the relevant background theories are already approximately _true_ [as well as instrumentally reliable]. On this assumption, the reliability of each of the principles in question with respect to the acceptance of _predictively reliable_ theories can be explained in terms of its contribution to the overall reliability of scientific practice with respect to the acceptance of theoretical principles and laws which are not only predictively reliable but approximately true as well.

Thus the first principle constrains us, _prima facie_, to accept only theories whose laws and ontologies closely resemble the laws and ontologies of theories already accepted. If those theories, in turn, provide a sufficiently accurate picture of the "furniture of the world" and how it works, then the operation of this principle will serve to make it more likely that theories which we take seriously are themselves approximately true.

Similarly, if well-confirmed theories are approximately true of real entities and if "measurement" and "detection" of theoretical entities really are measurement and detection, then there is no epistemological puzzle about the legitimacy of theoretical modification or extension of measurement procedures of the sort described. Adopting the detection procedures countenanced by such modifications is no more problematic epistemologically than applying the lens-makers equations to design a microscope and then using the microscope to observe bacteria.

Finally, the theory dependent principle of experimental design, if it operates against the background of a sufficiently accurate and suitably complete total science, will tend to isolate these respects in which a proposed theory is (speaking evidentially) most likely to fail if it is going to fail at all.

I propose (and I have argued elsewhere) that these explanations are, in fact, the only scientifically plausible explanations for the reliability of the theory-laden methodological principles in question, once it is remembered that the operation of these principles does epistemological work and that a purely pragmatic treatment of them ignores the vital questions about the instrumental reliability of scientific method, a reliability which even the most ardent contemporary empiricists do not question.

Thus Nagel's dictum is false: no adequate account of the logic and methods of science can be neutral with respect to the issue of realism; and this remains true even if one restricts one's concern to explaining those facts about the logic and methods of science which anti-realistic empiricists uniformly accept. What appears to be Smart's conception of the status of realism is likewise false: there is not one set of methodological principles appropriate for instrumental knowledge, and an additional set appropriate to the sort of theoretical knowledge which realists defend. Instead, the methods appropriate for theoretical knowledge are essential components of our instrumentally reliable methods.

2.2 Some Epistemological Lessons.

If the naturalistic and realistic account of the reliability of scientific method just sketched is approximately right, then several consequences follow which are significant for our understanding of the epistemology of science (and epistemology generally, for that matter). In the first place, theoretical considerations in science are evidential considerations. The methodological practice of preferring "simpler", that is, theoretically more plausible, theories contributes to the instrumental and theoretical reliability of the scientific method precisely because the fact that a proposed theory is theoretically plausible provides good reason to believe that it is (relevantly approximately) true. The theoretical plausibility of a theory constitutes genuine (if "indirect") evidence for the truth of that theory.

Moreover, the fact that theoretical considerations provide "indirect" evidence rather than "direct" experimental evidence is no indication that these are fundamentally different sorts of evidence. The assessment of "direct" experimental evidence is crucially dependent on theoretical considerations which reflect "indirect" evidential considerations. Experimental results are typically decisive in science, but their decisiveness depends upon prior (and evidential) judgments about the relation of the questions at issue to the theoretical tradition. Consistency with the results of observation and experiment is not the sole evidential criterion in the "experimental method".

We can put this same point in another way: What we have in the scientific method is a theoretical-presupposition-dependent total-science modification procedure, a procedure or strategy for deciding which modifications or additions to make to our existing body of accepted theories. <u>If</u> the total science with which we begin is relevantly sufficiently true and comprehensive, then the operation of this method will tend to ensure that later total sciences are successively more accurate and more comprehensive.

We may further explore the epistemological consequences of this conception of scientific knowledge by comparing and contrasting it with the epistemological positions represented by the traditional analysis of knowledge as justified true belief and by more recent causal theories of knowledge. For the purposes of this comparison, I will take as representative of received causal definitions of knowledge the claim that knowledge is reliably produced true belief [see Goldman 1967, 1976].

Both the traditional definition of knowledge and more recent causal theories are intended to provide an account of the nature of knowledge by providing a definition of knowledge which satisfactorily sorts cases of belief into cases of knowledge and cases of non-knowledge. What I want to argue in the remainder of this section is that--if the conception of scientific knowledge sketched in the preceding section is sound--then this strategy in epistemology is fundamentally unsound. It is not possible to draw the distinction between instances of knowledge and instances of non-knowledge in a philosophically revealing way (and this is so not merely because there will always be some "borderline cases"). Moreover, attempts to draw such a distinction are likely to obscure rather than to reveal certain absolutely fundamental features of knowledge. It will require some background work in exploring the similarities and differences between the realistic conception of knowledge sketched earlier and the more traditional accounts before a defense of this claim can be mounted.

Let us first consider the question in what sense the realistic theory of scientific knowledge of the first section is a causal or naturalistic theory. Two sorts of dissatisfaction with the traditional definition of knowledge seem to be the primary factors which have made causal or naturalistic theories seem attractive. In the first place, there are cases in which justification seems to be too weak a condition (in addition to belief and truth) for knowledge. There are cases in which one has an instance of justified true belief, but in which the justification at issue rests upon another belief which is defective in some way: it is false, or not itself known, or insufficiently justified.

A second class of cases shows that the requirement of justification may be too strong as well. These are cases in which the role assigned to a justification according to the traditional analysis is played instead by some fact of the matter which is relevant to the reliable production of belief, but not to justification as it is ordinarily understood. Here the clearest case is that of perceptual knowledge where the fact that the senses are reliable detectors plays the role which the traditional definition would assign to a justification.

It is pretty clear that the second class of cases is the more important for an understanding of the essential features of causal and naturalistic theories of knowledge. It is obvious how to try, at least, to modify the traditional definition to handle cases of the first sort. Cases of the second sort, if they really indicate that sometimes a brute fact plays the role assigned by the traditional definition to a justification, represent a far deeper challenge to the traditional conception of the nature of knowledge and the task of epistemology. To a good first approximation, we may characterize the traditional view this way: there are certain beliefs which have an epistemically privileged position; these might be beliefs about the content of sense experience, for example, or beliefs which have some appropriate sort of universality or innateness. At any rate, it is a matter to be decided _a priori_ what sorts of belief have this status. All instances of knowledge are either beliefs in this class or beliefs which follow from such privileged beliefs by appropriate principles of reasoning; here again, the significant point is that--although these principles of reasoning are not themselves always deductive--there are _a priori_ arguments which show that they are the right rules of "inductive" reasoning. The standards by which the justification for a given belief are to be assessed are themselves defensible _a priori_.

The new naturalistic or causal theories of knowledge depart from this traditional picture with respect to the role of _a priori_ principles in factual knowledge. In the first place, whatever epistemologically privileged status certain beliefs may have, their privileged status is a matter of contingent fact about the reliability of the relevant belief producing mechanisms rather than the result of _a priori_ considerations. More importantly, perhaps, the epistemic legitimacy of rules of inductive reasoning, or of inductive procedures generally, whether they involve reasoning or not, is not a matter of the _a priori_ justifiability of those strategies, but instead a matter of contingent fact about the reliability of those strategies in the actual world. This abandonment of a substantial part of the _a_ prioristic elements in traditional epistemology represents one of the two essential features of naturalistic epistemology.

The second essential feature of naturalism in epistemology is the unreduced appeal to causal notions in the analysis of knowledge. Largely because of the epistemological problems which surround the question of causal knowledge, traditional epistemology has treated causal notions as in need of (epistemologically motivated) explication rather than as ingredients in an analysis of knowledge itself. On epistemological grounds the empiricist tradition (which includes almost all contemporary English-language philosophy of science) has insisted on a non-realistic treatment of causal phenomena and causal powers: talk about causation is thought to be reducible to talk about regularities in nature. Naturalistic epistemology utterly breaks with this tradition: not only does it appeal to unreduced causal notions, despite the epistemological problems associated with causal knowledge, it employs such notions in the analysis of knowledge itself.

These two definitive features of naturalistic epistemology are, of course, closely related. Consider the question of the epistemological status of those features of our inductive strategy by which we discern regularities in nature. On the traditional account, the justification of these strategies is ultimately an a priori matter. Moreover, it is a constraint on empiricist analyses of knowledge and of causation that the a priori epistemological principles underlying the analysis of factual knowledge should explain why our procedures for obtaining knowledge of regularities in nature (which is all there is to causal knowledge on an empiricist view) are reliable or justifiable.

On the naturalistic view, by contrast, knowledge of regularities in nature need not be all there is to knowledge of causal relations and, moreover, an unreduced appeal to causal notions, rather than an appeal to a priori considerations, is required to explain the epistemological status of our strategies for discerning regularities in nature in the first place. Thus the two central features of naturalistic epistemology--the abandonment of a priori standards for the evaluation of inductive strategies and the employment of unreduced causal notions in the analysis of knowledge--are two sides of the same coin. We may, I believe, safely take that coin to be definitive of the naturalistic or causal approach to epistemology.

We are now in a position to see why the realistic account of scientific epistemology offered in the preceeding section is a naturalistic one. In the first place since, according to the realistic account, any evaluation of the evidence for a particular theory will depend upon prior theoretical commitments, the realistic account acknowledges the possibility that a scientific belief might be true and justified even though its justification rested upon background theories so thoroughly false that it should not be counted as an instance of knowledge. Thus, insofar as the insufficiency of justification for knowledge (given true belief) provides a reason for adopting a naturalistic epistemology, the realistic account provides a variety of examples of this phenomenon. Indeed, as we shall see, examples of this sort are much more significant for the naturalism of the realistic account than more ordinary examples of insufficiency of justification are for the defense of naturalistic epistemology generally.

Turning to the issues raised by the question of the necessity of justification for knowledge, we see at once that if naturalistic theories of perception, say, are correct then these theories will have to be incorporated into the realistic account of scientific knowledge insofar as that account addresses the role of observation in science. Much more importantly, however, the rest of the realistic account fully meets the criteria which we have seen as definitive of naturalistic conceptions of epistemology. The realistic account, like other naturalistic accounts, holds that the epistemic status of inductive strategies is not an _a priori_ matter. The actual inductive strategies which we employ at a given point in the history of science will reflect theoretical commitments characteristic of that time, _and_ these strategies will be reliable (even instrumentally reliable) only if the relevant theoretical commitments are nearly enough true and comprehensive. Since the truth and comprehensiveness of a body of scientific theories cannot be decided _a priori_, there are no _a priori_ standards sufficient for the epistemological assessment of actual scientific methods and practices.

Another way to put this same point is this: It is not the aim of the realistic account of scientific knowledge to deny that there are lots of cases in which justification is necessary for knowledge, if by justification one means, say, the justification of a particular experimental design by appeal to theoretical considerations. What the realistic account insists on is that the standards by which such justifications themselves are to be assessed are not (or, not wholly) _a priori_. The reliability of our practice of insisting on theoretical justifications of experimental designs depends on the approximate truth of relevant background theories, and _that_ is not a matter which can be determined _a priori_. When we treat the giving of justifications in science as a natural phenomenon, the question of its epistemic contribution to science is not an _a priori_ question.

The realistic account of scientific epistemology also satisfies the other criterion for naturalistic theories, the appeal to unreduced causal notions. Recall that if one says that knowledge is reliably produced true belief, then, in answer to questions about the epistemological status of inductive strategies, one will appeal to the unreduced causal notion of reliable belief production, rather than to purely _a priori_ epistemological principles. The realistic account of scientific epistemology is, if anything, more overt in its appeal to unreduced causal notions. It explains the causal reliability of scientific method with respect to instrumental knowledge not by appealing to _a priori_ principles but by assuming that the relevant background theories which determine the method's operation are approximately true and comprehensive descriptions _of the unobservable causal factors which underlie the relevant observable properties of_ _observable phenomena_. Whereas the empiricist tradition proposes to reduce talk of underlying causal powers or mechanisms to talk about regularities in nature, on the grounds that knowledge of underlying powers is impossible, the realistic account maintains that our knowledge of regularities in nature is parasitic upon our knowledge of underlying mechanisms.

Actually, the abandonment of _a priori_ considerations in favor of appeals to unreduced causal notions runs even deeper in the realistic account. Consider the fundamental principle of experimental design discussed in the preceding section. If the realistic account is right, this principle cannot be defended _a priori_ because its reliability depends on what one might call a "take-off point", a point in the development of the relevant scientific discipline at which the accepted background theories are sufficiently approximately true and comprehensive. When such a point has not yet been reached, the total science modification strategy, of which the fundamental principle of experimental design is a part, will not typically possess the sort of reliability which is characteristic of scientific knowledge. We cannot offer _a priori_ justification of the principle in question (or of the other principles characteristic of the total-science modification strategy) because we cannot show _a priori_ that a take-off point has been, or even will be, reached.

It might seem, however, that we could give a different sort of _a priori_ defense of these principles. We might, for example, be able to show _a priori_ that the fundamental principle of experimental design is a best possible principle for factual inquiry, and that no theory-independent principles which do not depend on take-off points are possible. But we cannot do this. We can certainly show that there are particular circumstances under which this principle is best possible (since our own circumstances are of this sort and we can show that they are). But it is easy to imagine possible worlds in which very simple and theory-independent principles of projection would suffice to obtain reliable instrumental knowledge, and in which efforts to obtain theoretical knowledge and to apply the principles which characterize _our_ scientific practice would result in utter failure. We can even imagine worlds in which nothing resembling experimental inquiry will be instrumentally reliable but in which instrumentally reliable laws are, say, revealed to those who pray in some appropriate way. Thus we cannot offer an _a priori_ defense of our take-off dependent principles as best possible (or even as good), nor can we find a methodology which _is_ defensible _a priori_ and which has our take-off dependent methodology as a special case.

Actually, what I have just said is not, strictly speaking, true. What is true is that issues about the epistemic reliability of methods or inductive strategies are _a posteriori_ issues, so that epistemology is one of the natural sciences, and methodological advances are, at least fundamentally, indistinguishable from advances in theoretical or practical knowledge. Since theories about the reliability of methods are ordinary scientific theories, we can, in fact, formulate a _meta-_

methodology which is take-off dependent in the sense that its reliability in guiding our methodological practice would depend upon the emergence of suitably approximately true and comprehensive *epistemological* theories. The maxim, "Use whatever methodology is best suited to obtaining true beliefs" is such a meta-methodological principle. Once a take-off point has been reached in epistemological theory, this maxim will dictate the adoption of more specific methodological principles which are themselves reliable. And, for all I can see, this meta-methodological principle might be defensible *a priori* provided it is understood as a take-off dependent principle.

But, there is no reason to believe that this meta-principle explains anything about the reliability of our actual scientific method. There is no reason to believe that we have reached a take-off point with respect to epistemology. Our methodological practices are theory-determined, and it is true that epistemological theories help to shape our methodological practices. But it is by no means clear that these particular theories have contributed positively to the reliability of scientific method. In the last generation of physicists, for example, a great many were persuaded of operationalism and other positivist epistemological doctrines. If the realistic account of scientific method is correct, these doctrines are profoundly false, and they probably made no positive contribution to the reliability of actual scientific practice. If we think of methodology as embodying discoveries about how the world works (in particular, about how approximately true beliefs can be obtained), then the development of the methodology of modern science must be seen as one of those cases in the history of science in which tacit knowledge far outstripped explicit knowledge. [It is worth remarking that such tacit epistemological knowledge is the sort of phenomenon which warrants Kuhn's insistence that paradigms in science amount to more than just the explicit theories they embody. [See Kuhn 1970.]]

At any rate, epistemology emerges, if the realistic account is correct, as the largely *a posteriori* study of a very complex natural phenomenon--the reliable development of successively more accurate and comprehensive theories and beliefs, about both observable and unobservable features of the world. It remains to see whether the complexity of the subject matter itself is compatible with the traditional project of seeking a definition of knowledge which will fruitfully sort cases of belief into cases of knowledge and cases of non-knowledge.

That scientific knowledge poses a problem for the philosopher who sets out to accomplish this task can be seen by considering two sorts of beliefs whose classification is especially troublesome. In each case, the trouble arises from more complex versions of the sort of situation which suggests that justification is not (given true belief) sufficient for knowledge, *viz*., cases in which true belief arises from beliefs which have, themselves, an epistemologically problematic status.

The first sort of belief consists of (approximately) true beliefs which arise early on in the development of some particular scientific discipline, roughly at the take-off point at which the relevant background theories in the field become sufficiently approximately true and comprehensive that a reliable methodology emerges. It would seem that, if we require that beliefs be reliably produced in order to count as knowledge, then the first generation of sufficiently true beliefs in any scientific discipline would not count as instances of knowledge. If, following somewhat more traditional theories of knowledge in this regard, we count beliefs as instances of knowledge only if they are grounded in beliefs which are themselves not only true but also known (or justified), then even the second generation of beliefs in an emerging science might be classified as non-knowledge. To make matters even more complicated, the question of the location of the take-off point itself may raise issues which will prove difficult to resolve. In the first place, there is some room for dispute about just what level of methodological reliability we should take as indicative of the emergence of scientific knowledge. More importantly, the reliability of method in a relatively new science need not be uniform: the relevant background theories may be accurate and comprehensive enough to serve as highly reliable guides in some investigations, while those same theories may prove less reliable in guiding other sorts of investigation. Along several dimensions the location of the take-off point may seem indeterminate.

The second class of cases includes those which reflect the "multi-perspectivicality" of knowledge attributions (Pastin 1978). Suppose that we are considering developments in a mature science whose take-off point is long past. Suppose that at sometime _t_ a Theory _T_ is proposed and is subjected to all the right sorts of experimental test, that is, all the sorts of tests which the fundamental principle of experimental design dictates, given the background theories accepted at _t_: Even in a mature science it could happen that at some substantially later time _t'_, new theoretical discoveries are made whose effect is to suggest an alternative to _T_ which could not have been anticipated at _t_. The fundamental principle of experimental design, applied at _t'_, will require that _T_ be tested against this new alternative. Suppose that it is so tested, and that it remains well confirmed. Suppose also that it is in fact true. Now, imagine a group of specialists in the field talking about the old days, and considering the question of whether, _at t_, scientists "really knew" that _T_. One side maintains that the earlier scientists certainly did know that _T_ since _T_ was true and since their methodology was impeccable. The other side replies that the earlier scientists' experimental evidence for _T_ would not be acceptable by current standards because of their (quite non-culpable) failure to exclude the more recently considered alternative to _T_, and therefore these earlier scientists did not "really know" that _T_.

The two sides evaluate the methodology of the earlier scientists from different perspectives. Each side correctly applies its own perspective. But it may well seem indeterminate which of these two perspectives, if either, is appropriate to answer the question at issue. Moreover, it certainly seems clear that, whichever side we might take to have won the argument, the other side is on to an _epistemologically_ important fact about the history of the theory at issue.

It is perhaps worth remarking that there are interesting cases which are intermediate between cases involving multiperspectivicality and those which reflect questions about the location of the take-off point. I mentioned, earlier, cases in which the relevant background theories in a discipline might be sufficiently accurate and comprehensive to reliably guide some investigations, but might be less reliable with respect to others. It is also possible that within a single investigation (or class of investigations) the relevant background theories may reliably guide some features of the investigations (like, say, the identification of the significant causal factors in some sorts of natural phenomena) but be inadequate to guide other features of the same investigations (like, say, elucidation of the details of the reaction mechanisms involving the causal factors at issue). In cases such as these, our overall assessment of the reliability of the belief producing mechanisms in the discipline will depend on which features of the research in the discipline we wish to emphasize. Here again there will be cases in which it seems indeterminate which features are most significant if one wishes to decide whether the belief producing mechanisms _are_ sufficiently reliable to produce knowledge.

It remains to show that these cases of apparent indeterminacy are cases of indeterminacy, that is, to show that it would be a mistake to look for theories of knowledge which will satisfactorily resolve them one way or another. If I am right, theories of knowledge which attempt such a resolution would have to either ignore or underestimate the importance of the dialectical elements in scientific knowledge. By way of seeing why this is so, consider what modifications in the formulation of causal theories of knowledge are dictated by the naturalistic and realistic account of scientific knowledge which we have been discussing.

In the first place, it is clear that the issue in scientific knowing is not the reliable _production_ of beliefs. The reliability which scientific practice displays is not so much a matter of how beliefs are produced or even of how they are initially accepted, but of the tendency over time for beliefs to be sustained only if they are approximately true _and_ for beliefs to be modified in the direction of closer and closer approximations to the truth. What is at stake is reliability in the _regulation_ of belief (over time) rather than reliability in the _initial_ production _or_ _acceptance_ of particular beliefs. Indeed, I think that this will prove to be true in many cases of everyday belief as well.

Secondly, the notion of _exact_ truth plays no significant role in the realistic account of the reliability of scientific methodology. The reliability of the scientific method does not depend on the exact truth of background theories, nor does the operation of that method typically produce beliefs which are, strictly speaking, exactly true. Indeed, there is considerable evidence for the truism that we know now that all the theories we accept are in some respects false. Exactly true theories, if there have been any at all, are utterly exceptional in science (and, indeed, in those areas of everyday knowledge in which statements of great precision are made).

It follows, therefore, that if by "knowledge" we refer to the sort of thing which careful everyday and scientific investigations aspire to and sometimes achieve (as the theory of reference defended later in this essay suggests), then (exact) truth is not necessary for knowledge. One *might* argue that truth *is* necessary for knowledge, and that corresponding to every case of (say) reliably regulated approximately true belief, there is the real knowledge that the belief in question is approximately true. But since, in the whole naturalistic account of the cases which we ordinarily count as knowledge, no reference to exact truth plays *any* explanatory role, it is difficult to see why a naturalistic or causal theory of knowledge should treat (exact) truth as necessary in this way. Not even in a usefully idealized conception of scientific knowledge will truth emerge as a necessary condition. Indeed an idealization which portrays exact truth as necessary for knowledge would obscure the deepest epistemological facts about scientific inquiry.

These last two points are really both reflections of the dialectical character of the scientific method (and of everyday reasoning for that matter). It is mistaken to define knowledge in terms of exact truth *or* in terms of reliable belief *production* (or justification *at a time*) precisely because the natural phenomenon in which knowledge is manifested involves a dialectical process of *successive approximation* to the truth, whose reliability consists in a tendency *over time* for the successive approximations to be increasingly accurate. Both the traditional definition of knowledge and the version of the causal theory we are examining are inadequate because each presents a picture of knowledge which is static rather than dialectical.

How, then, does a dialectical and naturalistic account of knowledge help us to understand the question of the classification of the first or second generation of (approximately) true theories within a given scientific discipline, or the question of classification which arises from multi-perspectivicality? The surprising answer, I believe, is that the lesson we should draw from examining the naturalistic and dialectical aspects of scientific knowledge is that it should not be the aim of a theory of knowledge to effect such a classification at all. If the aim of epistemology is to say what knowledge really is, then it should not be part of the aim of epistemology to resolve such disputes about the boundary between those beliefs which are instances of knowledge and those which are not. Epistemology should abandon this project not primarily because there are lots of tricky "borderline cases" of knowledge, but rather because, if knowledge is a dialectical matter, then the project is basically misconceived.

Several considerations seem to me to dictate this conclusion. To begin with, we already know what knowing is when it is considered as a natural process: it is reliably regulated believing. If we ask how a particular belief might relate to the processes of reliable belief regulation, several points become clear. In the first place, if we are concerned with the way in which these processes impinge on the belief in question (its acceptance, rejection, or modification), then there will not, in general, be a time at which the possibility of future acceptance, rejection, or modification is finally excluded. Our assessment of the evidence for or against a theory or other belief will always reflect the particular background beliefs which are themselves accepted at the time the assessment is made. Since these beliefs change over time, our evidential standards change as well.

It is just this sort of change in standards which gives rise to the puzzle of multiperspectivicality. This very dialectical evolution of evidential standards is central to the reliability of the scientific method. As successively more accurate theories are accepted, our evidential standards become successively more reliable, thus facilitating the adoption of more accurate theories, and so on. For any particular belief, and any particular time, we can "freeze" this process if we want to, and inquire how the belief fares with respect to the evidential standards _at that time_. But the answer to this question will not, in general, be epistemologically significant. Or, rather, the answer may be quite significant (if, for example, the time of the "freeze" is the present), but the answer will be of no help if our aim is to resolve the problems of belief classification raised by the issue of multiperspectivicality. For a given belief there is, in general, no special point in the history of science such that the evidential assessment of the belief _at that point_ (or after it) is _the_ epistemologically definitive judgment. Nor is there any ahistorical stance from which an epistemologically privileged assessment can be made. From an epistemological point of view there is nothing more to say than to lay out the facts which pose the presumed dilemma. Indeed, the fact that--in such cases--there is nothing more to say is itself a reflection of the dialectical and cumulative features of scientific method upon which its reliability depends.

Thus, we are justified in treating such cases of multiperspectivicality as genuine cases of indeterminacy. To insist that there is some epistemological fact of the matter which settles the "dilemma" one way or another would not be merely to insist on the reality of a distinction which would in fact have to be drawn arbitrarily. It would be to treat as puzzling a phenomenon which, from the point of view of naturalistic and dialectical epistemology, is entirely straightforward; to treat as a dilemma, a question which is in fact misconceived.

A similar situation obtains with respect to the question of the identification of the take-off point in the history of particular sciences, and the question of classification raised by the first and second generation theories in emerging sciences. Consider the question of when the methodology of a particular science first attains the reliability characteristic of scientific knowledge. Scientific method constitutes a total science modification strategy; moreover, since the method itself is theory-dependent, it embodies a procedure for its own modification. The epistemic reliability of the scientific method at a time is manifested not only its reliability as a total science modification strategy, but also in its (consequent) reliability as a methodology modification strategy.

There are thus two important reasons for treating the question of the location of the take-off point in a science as indeterminate. The problem is not that there would have to be some degree of arbitrariness in any decision about the location of the take-off point. The problem is that--along two related dimensions--scientific method is less static than the question at hand suggests. In the first place, there is something misleading about asking about the reliability of the scientific method at a particular time: the reliability characteristic of the successful operation of the scientific method is displayed over time, in the operation of a dialectical process of total science improvement. One can ask, at a particular time and for a particular scientific field, how successful the scientific method is at that time in guiding the acceptance of new or modified theories, but such a static assessment of the reliability of scientific methodology will not, generally speaking, reflect an accurate assessment of its longer-run tendency to produce increasingly accurate theories.

The fact that the scientific method also functions as a methodology improvement strategy adds an additional dimension of indeterminacy to the question of when a particular field's methodology becomes reliable enough to count as fully scientific. Part of the reliability of methodology within a field may consist in the extent to which the theoretical developments countenanced by that methodology lead to improvements in the methodology itself. In this case too, the static question of how reliable the methodology of a field is, at a particular time, with respect to the production of methodological progress, is misconceived: the sort of reliability in question, when it obtains, is manifested in a dialectical development over time. Thus, just as in cases of multiperspectivicality, the epistemologically relevant facts about the emergence of the scientific method within a discipline are exhausted by a recitation of those facts which, seen from a non-dialectical perspective, would seem to pose the challenging question of when the discipline became fully scientific. Here too, the issue is indeterminate not because any answer to the question would have some aspects of arbitrariness, but rather because the very question presupposes a mistakenly static conception of knowledge.

With respect to the question of whether the first or second generation of approximately true theories within a scientific discipline are instances of knowledge, the situation is even more favorable to a diagnosis of indeterminacy. We have already seen that the relevant epistemic relation between a scientific theory and methodological practice is a continuing and dialectical one, and that there is a misconception in the view that there will always be a cogent answer (even a partly arbitrary answer) to the question of whether a theory or other belief is an instance of knowledge at a particular time. We have seen, moreover, that this sort of indeterminacy is best exemplified when there are significant shifts in theoretical understanding of the sort which is certainly characteristic of the earliest stages in the emergence of a new scientific discipline. We have also seen that there is an indeterminacy involved in the question of when such disciplinary emergence takes place.

These considerations no doubt suffice to show that it should not be the aim of epistemology to answer the questions of classification at issue. But there are further considerations which also dictate this conclusion. In the first place, there are certain theoretical developments which occur early on in the history of a scientific discipline and which are essential to the establishment of reliable methodology. The acceptance of the atomic theory of matter, say, or of Newtonian mechanics was crucial in the evolution of reliable methodology in chemistry and physics. When the acceptance of a theory is thus _a constituent_ of the establishment of reliable methodology, the question of its relation to reliable methodology is even more dialectically complex than the same question asked about theories whose adoption was less crucial to the establishment of a reliable "normal science" (see Kuhn 1970, chapters 1-5; the naturalistic interpretation of the emergence of normal science offered here is, of course, quite different from Kuhn's interpretation). Moreover, it is typically the case that the subsequent rigorous evidential evaluation and (re-?) confirmation of the theories which thus establish a reliable research paradigm depend upon theoretical and experimental developments which are made possible only by the theories' original adoption. The best continuing evidence we have for theories of substantial scope often arises from the continued successful articulation of the research paradigms of which they form the basis. (Here, again, the terminology is Kuhn's, but the interpretation is non-relativistic; for a further discussion of this approach to the notion of paradigm establishment and paradigm articulation see Boyd 1979.) Finally, if the realistic account of scientific methodology is correct, the contribution which the acceptance of a theory makes to the establishment of a reliable methodology itself provides some of the evidence for the approximate truth of the theory--and this is true not only for the earliest and most general theories within a paradigm, but for theories generally.

All of these considerations indicate that, in the case of questions about the epistemic status of early scientific theories, as in cases of multiperspectivicality, the epistemologically relevant facts are _just_ those which make the issue of whether the relevant belief is an instance of knowledge or not seem so intractable. No further epistemological considerations can resolve such an issue; the complex and dialectical story is all there is to say. It does not follow, of course, that we can _never_ say, of a belief (scientific or otherwise) and a time, that the belief is definitely an instance of knowledge at that time, or that it definitely is not. Such classifications are often possible. But the standards for such classifications do not reflect the most important features of scientific knowledge; I would be quite surprised if the same were not true for interesting and complex non-scientific knowledge--moral knowledge, for example, or intuitive knowledge of human psychology. At any rate, if the dialectical and realistic conception of scientific knowledge is correct, then the fundamental aim of epistemology should be the naturalistic elucidation of the mechanisms of reliable belief regulation rather than the formulation of a general definition of knowledge, considered as a property of individual beliefs.

Consider an actual case in the history of science: Suppose someone asks whether 17th-century mechanists "knew" that the corpuscular theory of matter which they advocated was (approximately) true, or whether the 16th- and 17th-century founders of modern physics and chemistry invented "the scientific method" as we now know it. If the epistemological perspective which I have articulated here is correct, we should answer roughly as follows:

The scientific method is theory-dependent and it regulates believing with the reliability we think of as typical <u>only</u> when the relevant background theories are sufficiently complete and sufficiently accurate. The reliable regulation of believing typical of scientific practice depends not only on our employing a rigorous method, but also upon our having approximately true theories to start with. Before that happens, nothing like what we think of as modern science is possible.

Thus, the triumph of the scientific revolution cannot be a triumph of a method which excludes presuppositions. Furthermore, the triumph of the scientific revolution can't be a triumph of method <u>alone</u>, anyway. What makes the 16th- and 17th- centuries central in science is not just a change in the importance accorded to experiment and observation, but also the fact that atomistic and mechanistic theories emerged and that they happened to be relevantly approximately true.

We cannot say the mechanism of Descartes, Newton or Boyle was dictated by rational scientific method <u>as we now know it</u>--because the relevant approximately true background theories were largely absent until such scientists as these proposed them. What we call the scientific method didn't have its truth-generating capacity until scientists hit upon enough relevantly correct theories.

This is not to say that 17th-century mechanists were irrational--their guess was a natural extension of laws governing obviously mechanical phenomena. But neither were "Renaissance naturalism" in chemistry and talk about the "dormative powers" of opium <u>then</u> irrational.

Here then is a case in which it is reasonable to doubt whether epistemology should sort knowledge from non-knowledge:

Boyle's corpuscularism was in some respects true.

It was central in establishing the possibility of scientific knowing in chemistry.

But was Boyle's theory itself reliably produced? Well, in a sense, no, since scientific reliability in chemical methodology was only made possible by its adoption.

On the other hand, in a sense, yes, since mechanistic suspicions of "occult powers" helped to produce it--and such suspicions in fact constituted a reliable belief regulating principle.

But were these suspicions reliably produced? Well, in a sense...

We are right--in a way--to say, as we do, "Boyle knew that matter was composed of small particles, which he called 'corpuscles'."

We are also right--in a different way--to say, as philosophers sometime do, "Boyle didn't really know that the atomic theory of matter was true."

Should epistemology decide between these alternatives? Almost certainly not. "Knowledge" is a relatively non-specific honorific--there are clear cases of beliefs which we warrantedly honor with it and clear cases where we deny the honor. For the interesting cases in between, the naturalistic epistemologist will simply tell the tale of the dialectical relation of the beliefs in question to such reliable or unreliable features of belief regulation as are relevant. There will be nothing more to say. It will not generally even be possible to treat the classification of approximately true beliefs as knowledge or non-knowledge as a matter of degree. The features of scientific practice which make it epistemologically reliable (when it is reliable) are quite diverse: approximate truth of background theories, soundness of experimental design, emphasis on observational/experimental method, appropriateness of metaphysical "hunches" (like the anti-"animistic" hunches of 17th-century mechanists), freedom from prejudicial political interference, reliability of the indoctrination of graduate students with respect to the more "intuitive" and as yet unarticulated features of the "paradigm", etc. Although these features are all related--logically and epistemologically and causally--they do not represent anything like a single dimension along which the reliability of belief regulation can be assessed. The problem of classifying beliefs as knowledge or non-knowledge is not merely that there are border-line cases. The problem, instead, is that the reliable regulation of belief has too many important dimensions, interacting in too complex a way.

Consider the question--"Are very long polymer chains single molecules?" The sophisticated chemist will explain the various ways in which the micro-structure of polymers is similar to, and different from, the micro-structure of typical molecules. She will not get caught in a pointless discussion about whether to withhold or apply the term "molecule". The micro-structural story is all there is to say.

Mechanisms of reliable belief regulation are the micro-structure of knowing. Like polymer chemists, naturalistic epistemologists will often find that a micro-structural story is all there is to say.

One remaining point must be made regarding the conception of knowledge which emerges from the realistic account of scientific epistemology. Logical empiricists maintain that all factual knowledge is empirical knowledge, that all factual knowledge is grounded in the results of observation and experiment. Empiricists deny that any factual statements can be known *a priori*; they deny that there are any factual presuppositions of our knowledge which are immune, in principle, from refutation by experimental evidence. *In no respect whatsoever does the realist conception of scientific knowledge offered here challenge these doctrines*.

Where realist and empiricist accounts of knowledge differ is in their understanding of the nature of experimental evidence and "inductive" inference. The empiricist tradition takes the epistemic primacy of observation and experiment to entail that evidence for a factual hypothesis must (when it exists at all) consist only in the confirmation by observation or experiment of certain of the observational consequences of the hypothesis at issue (together with "auxiliary hypotheses"). The realist conception, by contrast, asserts that:

(1) Theoretical considerations in science are also evidential: judgments of the "simplicity" of a theory or of the support it receives in virtue of its theoretical plausibility (which are, roughly, the same thing), are judgments about the evidence for or against the theory.

(2) There are no theory-neutral standards of "direct" experimental evidence for a theory. The evidential support (or potential disconfirmation) which a theory receives from tests of its observational consequences depends crucially on theory-mediated evidential considerations.

(3) The reliable operation of the scientific method depends on the (contingent) emergence in various scientific disciplines of suitably approximately true theories. After the emergence of such theories, the scientific method displays its characteristic epistemic reliability. It functions as a theory-dependent total-science *improvement* strategy whose operation produces additions to, revisions of, and, often, outright disconfirmation of, the particular laws and theories which form our total science at any particular time.

This conception of factual knowledge is fundamentally different from the empiricist conception, especially in its emphasis on the role of contingent factors in the reliability of scientific method, and in its dialectical emphasis. But it does not deny that all factual knowledge is empirical or experimental knowledge. It portrays all scientific knowledge as grounded in observation and experiment. It does not countenance the existence of *a priori* factual knowledge. It does not say that the presuppositions of scientific research are immune from refutation. Instead, it provides an explanation of the very method by which we subject our theoretical and instrumental conceptions to rigorous experimental test.

Insofar as the attractiveness of the empiricist conception of knowledge rests upon the conviction that factual knowledge must be empirical knowledge, the realistic conception will, in that respect, be equally attractive. Their dispute is not over the primacy of experimental knowledge, but over its nature.

Similarly, it might seem that the realistic methodology advanced here is less "rigorous" than that which would be countenanced by logical empiricism. The empiricist accepts only "direct" experimental evidence as relevant to the acceptance or rejection of a proposed theory; that is, only such evidence as is reflected in the confirmation or disconfirmation of one of the theory's observational predictions. The realist, on the other hand, countenances theoretical considerations as providing evidence relevant to theory confirmation: evidence which reflects a theoretical assessment of the plausibility of a theory in the light of previously accepted theories. Even if this sort of theory-mediated evidence is portrayed as "indirect" experimental evidence, since it rests upon background theories which have themselves been experimentally tested, it might seem that its acceptance would reflect the adoption of less rigorous standards. After all, the empiricist counts _only_ its successful observational predictions as evidence for a theory, whereas the realist might also count the theory's theoretical plausibility as _additional_ confirmatory evidence. Countenancing this additional source of confirming evidence will inevitably result in less rigorous standards of evidence.

This conception of the issue is mistaken. The issue between realists and empiricists is not over the appropriate degree of evidential rigor, but instead one over the nature of such rigor. In the first place, theoretical considerations can count for the disconfirmation of a theory as well as for its confirmation, so that it is by no means even _prima facie_ plausible that countenancing theoretical considerations as evidentially relevant would make it easier for a theory to be confirmed. More importantly, if the realistic account of scientific method presented here is correct, then "indirect" or theory-mediated evidential considerations are absolutely essential to the _rigorous_ assessment of the evidence for or against a proposed theory. This is true because theory-mediated evidential considerations are essential in identifying the experimental conditions under which a proposed theory must be tested, in order that a rigorous assessment be possible. The fundamental principle of experimental design requires precisely that we test a proposed theory against alternatives which are themselves supported by "indirect" or theory-mediated evidential considerations. A methodology which did not countenance such "indirect" evidential considerations would be one in which rigorous experimental testing of theories would be impossible.

Analogous considerations permit a rebuttal to a closely related unfavorable comparison of realism with logical empiricism. It is plausible to hold that empiricists are more rigorous (or, at any rate, more cautious) than realists just because, given any body of experimental evidence, the realist will be inclined to count more beliefs as confirmed than will the empiricist (since each might take the instrumental reliability of certain theories to have been confirmed, and the realist would take their approximate truth to have been confirmed as well). On the contrary, if the realistic and dialectical conception of scientific method is correct, then the empiricist who does not accept theoretical knowledge would be disbarred from employing the very features of the scientific method upon which empirical rigor depends.

Of course, an empiricist might simply adopt the theory-dependent and dialectical method of science, while denying that theoretical principles in science embody knowledge of the world. Such an empiricist would, if what I have said here is correct, be employing a method whose reliable operation (s)he could not explain. The risk of false theoretical belief would thereby be eliminated at the cost of the employment of a method which the empiricist would not and could not fully understand. There would be, for such an empiricist, the very serious risk of error with respect to crucial methodological issues whose successful resolution would require an accurate understanding of the way in which the scientific method actually works. Thus, the empiricist we are envisioning would run a long term risk of error with respect to instrumental knowledge, arising from a defective understanding of the scientific method itself. The suggestion that such a long term risk would arise from rejection of realistic methodological principles is hardly speculative, as the history of behaviorism in American psychology testifies.

Finally, it is by no means clear that, even in the short run, the peculiar empiricist we are considering would improve his/her ratio of approximately true beliefs to significantly false ones. The scientific method is not, generally speaking, more reliable at producing instrumental knowledge than at producing theoretical knowledge. Depending on the particular discipline, and the particular historical period, the empiricist we are considering might end up with an account of nature which was, on balance, less accurate as well as less complete than that accepted by his/her more realistic counterpart.

Thus scientific realism differs from empiricism neither in the extent of its commitments to the experimental method, nor in the rigor of its evidential standards, nor in the degree of inductive caution it recommends. Instead, the dispute is over the nature of the experimental method, of evidential rigor, and of inductive caution.

2.3. Natural Kinds and Scientific Knowledge

According to traditional empiricist accounts, the kinds into which we classify natural phenomena are largely arbitrary. Our categories do not "cut the world at its joints" or sort things according to their "real essences". Instead, the boundaries of our categories are determined by arbitrary definitional conventions and, perhaps, by innate standards of similarity which reflect the structure of the human mind rather than the structure of external reality. The traditional (and the contemporary) arguments for this position are largely epistemological: the unobservability of "real essences" or "joints" in nature is seen as precluding knowledgeable classification of natural phenomena into the categories which they allegedly define.

The decline of verificationism has been associated with the emergence of more realistic and naturalistic treatments of the issue of kinds and classification. These treatments have fallen roughly into two groups. Some treatments (e.g., Putnam 1975) have emphasized that we sometimes classify things according to what might be called their explanatory real essences: the fundamental (and, perhaps, unobservable) properties or underlying mechanisms which must be invoked in successful explanation of their observable properties or causal powers. Other treatments (e.g., Quine 1969, chapter 5) hold that the "naturalness" of some kinds consists in their appropriateness for successful inductive generalization. The realistic and naturalistic conception of knowledge presented in the previous sections provides the basis for a fruitful integration of these two perspectives. The integration is, in fact, anticipated very early in the empiricist tradition.

Locke, in Book Four of the Essay Concerning Human Understanding, reaches the striking conclusion that general knowledge of the sensible properties of bodies is impossible without knowledge of their insensible corpuscular structure.

"...But whilst we are destitute of senses acute enough to discover the minute particles of bodies, and to give us ideas of their mechanical affectations, we must be content to be ignorant of their properties and ways of operation; nor can we be assured about them any further than some few trials we make are able to reach. But whether they will succeed again another time, we cannot be certain. This hinders our certain knowledge of universal truths concerning material bodies: and our reason carries us herein very little beyond particular matters of fact." (Locke, Essay, IV, iii, 25).

(See also IV, iii, 14; IV, iii, 29; IV, vi, 131. Also Hume, Treatise, Book I, Chapter III, sections IV and XIII; Hume, Inquiry, Section IV, part 2, Section V, part 2).

Part of the remedy to Locke's pessimism regarding general knowledge of substances is provided (probably tongue-in-cheek) by Hume and, in a more serious and sophisticated treatment, by Quine:

> Here, then, is a kind of pre-established harmony between the course of nature and the succession of our ideas; and though the powers and forces by which the former is governed be wholly unknown to us yet our thoughts and conceptions have still, we find, gone on in the same train with other works of

nature... .(Hume, Inquiry, Section V, part II).

One part of the problem of induction, the part that asks why there should be regularities in nature at all, can, I think, be dismissed. That there are or have been regularities, for whatever reason, is an established fact of science; and we cannot ask better than that. Why there have been regularities is an obscure question, for it is hard to see what would count as an answer. What does make clear sense is this other part of the problem of induction: why does our innate subjective spacing of qualities accord so well with the functionally relevant groupings in nature as to make our inductions tend to come out right? Why should our subjective spacing of qualities have a special purpose in nature and a lien on the future?

There is some encouragement in Darwin. If people's innate spacing of qualities is a gene-linked trait, then the spacing that has made for the most successful inductions will have tended to predominate through natural selection. Creatures inadvertently wrong in their inductions have a pathetic but praiseworthy tendency to die before reproducing their kind.

. . .

He [man] has risen above it [his innate subjective spacing of qualities] by developing modified systems of kinds, hence modified similarity standards for scientific purposes. By the trial-and-error process of theorizing he has re-grouped things into new kinds which prove to lend themselves to many inductions better than the old.

. . .

A theoretical kind need not be a modification of an intuitive one. It may issue from theory full-blown, without antecedents; for instance, the kind which comprises positively charged particles.

We revise our standards of similarity or of natural kinds on the strength, as Goodman remarks, of second order inductions. New groupings, hypothetically adopted at the suggestion of a growing theory, prove favorable to inductions and so become 'entrenched'. We establish the projectibility of some predicate, to our satisfaction, by trying to project it. In induction, nothing succeeds like success. (Quine 1970, pp. 126-129).

Although such naturalistic replies as Quine's allow us to avoid Locke's extreme scepticism regarding general knowledge of bodies, there are, nevertheless, two important grains of truth in Locke's treatment of the issue of inductive generalization. First, if our generalizations regarding the sensible properties of bodies are to be reliable, then there must--often enough--be an important nonarbitrariness about the categories which we employ in formulating our generalizations: these categories must correspond in some appropriate way to whatever the relevant causally determining features of the microstructure (or functional arrangements, etc.) of bodies are. The point of Quine's appeal to natural and to cultural evolution is to indicate that a correspondence of this sort need not be mysterious or magical even when it does not arise from knowledge of the relevant underlying features.

If the account of the epistemology of science offered in the preceding sections is correct, then the second, and more surprising, grain of truth in Locke's treatment of induction is that knowledge of determining microstructural properties of matter *is* required for successful inductive generalization about the sensible properties of matter, *if* such generalizations are to have the scope and precision typical of mature sciences. In particular, Locke was right to think that successful inductive generalization in chemistry would depend crucially upon further knowledge of the "corpuscular" structure of chemical substances (though wrong, of course, to think that such knowledge would require more acute senses). In general, classification of sensible things on the basis of knowledge of their unobservable properties *is* a prerequisite for successful and sophisticated inductive generalization regarding their observable properties.

Here then is the basis for an integration of the two naturalistic conceptions of kinds. Kinds characterized by "explanatory essences" are also kinds from the point of view of inductive generalization; indeed, in mature sciences, kinds which are explicitly characterized in terms of explanatory essences are the overwhelmingly typical cases of inductively natural kinds. Kinds natural from the point of view of successful induction need not always be explanatorily natural kinds, but they must correspond in relevant respects to the (perhaps unobservable) properties and mechanisms which causally determine the observable properties of the subjects of empirical generalizations. Moreover, an understanding of such determining properties and mechanisms is the procedure for identifying inductively natural kinds which is characteristic of scientific inquiry. The social and cultural evolution of our inductive categories proceeds (especially since the 17th-century) by the discovery of natural kinds in the explanatory sense [see Boyd 1979].

If this naturalistic understanding of kinds and categories is right, then several interesting consequences follow. In the first place, it is clear on either conception of kinds, but especially clear given the integration of these conceptions, that "natural" kinds are relative to disciplines, inductive tasks or contexts of inquiry. A scheme of classification which is inductively appropriate and micro-explanatory with respect to one set of questions, or one sort of inquiry may be inappropriate in another context in which, say, different aspects of underlying micro-structure (or functional organization, or ecological setting, ...) are causally relevant to the factors under study. Thus, in a certain sense, human interests, projects and practices are partly definitive of natural kinds. But this sort of dependence of kinds upon human interests and human inquiry is fully compatible with the realist idea that reality is prior to thought and prior to human action in general. No implication of idealism or of social construction of reality follows. Instead, what is indicated is the complexity of the causal structure of the world to which we must accommodate our intellectual and social practices.

The theory-dependence of those schemes of classification which are either inductively or explanatorily fruitful also helps to explain why, as Quine insists, there are no (or, perhaps very few) terms in natural languages which possess analytic definitions. Such terms would correspond to kinds or categories whose boundaries would be in an important sense arbitrary, since they could not be revised in the light of further evidence or theoretical discoveries. But it would be sheer luck if a conceptual scheme in which such categories predominated would be useful either for inductive generalization or for explanation. Thus there are very strong reasons to believe that only the most intellectually trivial terms in a successful language could possess unrevisable definitions.

In a similar fashion, the naturalistic account of kinds allows us to see the defects of a certain conception of philosophy as "rational reconstruction". One of the standard research strategies in philosophy is to seek to discover--for certain important concepts--necessary and sufficient conditions for membership in the class which they define. One justification for such a strategy is that such conditions will be uncovered when the relevant analytic definitions are made explicit. But even philosophers who are uncomfortable with the analytic-synthetic distinction may hold that it is the task of philosophy to rationally reconstruct our conceptual schemes so that such conditions can be articulated and borderline cases resolved. Even if this is a partly empirical enterprise, it might be argued, still it is part of the task of philosophy to clarify and refine our concepts in this way.

Against this quite plausible view, it may be argued that there will be cases in which the "fuzziness" of our (pre-philosophical) concepts reflects, not the lack of suitable philosophical subtlety, but the actual complexity and multi-dimensionality of the underlying causal structures upon which the "naturalness" (explanatory or inductive) of our categories depends. In such cases, boundary-resolving rational reconstruction would obscure, rather than clarify, the "essences" of the kinds in question. It is precisely this sort of situation which, I have argued, obtains in the case of the concept of knowledge. Thus a naturalistic treatment of natural kinds and categories vindicates some of the more peculiar features of the naturalistic account of knowledge offered earlier in this essay.

Finally, consideration of Quine's treatment of natural kinds affords us a significant example of an important tension in recent non-positivist philosophy of science. Philosophers who have rejected verificationism and who are sympathetic to some version or other of scientific realism often subscribe to a conception of the epistemology of science which conforms to what I have been calling Nagel's dictum. In their treatment of methodological issues they are not nearly so straightforwardly realist as they are with respect to ontological issues. Now, Quine is, in some perfectly straightforward sense, a scientific realist--a willingness to accept the theoretical posits of current physical theory is almost a hallmark of his views about science. Indeed, in the essay we are discussing he gives the example of the class of (presumably unobservable) positively charged particles as the example of a theoretical kind.

Nevertheless, when Quine describes the method by which theoretical kinds are discovered, he follows Goodman in describing the process as "second-order induction". To be sure, this is right--scientists do make second-order inductions about which categories and kinds are appropriate for (first-order) induction. But, as Quine recognizes, the way in which they identify those theoretical kinds is by paying attention to theoretical considerations. If the account offered here is correct, the theoretical inferences which they make are, in the first instance, first-order inductions about the causal properties of theoretical (and observable) entities. The reliability of their second-order inductions is almost entirely parasitic upon the reliability of these first-order theoretical inductions. But to say this is to abandon Nagel's dictum and to break fundamentally with the empiricist tradition on issues in the epistemology of science. Quine's preference for the less overtly realist characterization of these inferences as second-order inductions illustrates, I believe, the residual effect of empiricist doctrines in epistemology on the work of scientific realists. If the conception of scientific epistemology defended here is correct, this residual influence has the effect of partly obscuring some of the most important realist insights into the nature of scientific methodology.[2]

2.4. On Reference, Univocality and "Unity of Science"

The second methodological principle discussed in 2.1 is, in fact, a special case of the methodological principle which logical positivists called "unity of science". According to this principle, any number of different well-confirmed theories [perhaps from quite disparate scientific disciplines] may be jointly employed with the expectation that the observational predictions deduced from them will be accurate. Unity of science is an epistemological principle: it says something about which inferences regarding observational predictions are justifiable or reliable. It is also an epistemologically puzzling principle, especially for philosophers with an anti-metaphysical bent. Indeed, 20th-century positivists often offered it as part of a "rational reconstruction" of metaphysical materialism.

To see what the puzzle is all about, consider the following special case of "unity of science" (special only because it refers to only two theories):

If T' and T'' are any two well-confirmed theories, from whatever scientific disciplines, then--even if no observational test has yet been made of a prediction deduced from their conjunction--it is reasonable to expect that predictions deduced from their conjunction will be (approximately) true.

Here we can see a puzzle: The predictive reliability of $(T' \wedge T'')$ can apparently be assessed even though none of its observational predictions whatsoever have been subjected to observational or experimental test.

In order to see the issue raised by this puzzle more clearly--and, in order to get the principle of unity of science formulated right--it is necessary to recognize that the principle, as stated, requires two modifications.

(i) In the first place, there are some circumstances in which we are in a position to be pretty sure that the conjunction of two well-established theories, T' and T'', will <u>not</u> be predictively reliable. Such cases occur whenever we have good reason to believe that T' and T'' are each approximations, applicable in different circumstances, to some more general theory which has not yet been formulated. The principle of unity of science needs to be amended to exclude such cases.

More important for our purposes here is the following:

(ii) In stating "unity of science" we must recognize explicitly something which is taken for granted when the principle is applied in practice, <u>viz</u>., that each of the theoretical terms (and observational terms too, for that matter) which occur in T' and T'' should be <u>univocal</u> in its occurrences in their conjunct.

That such a condition on applications of the unity of science principle is necessary must be obvious; without it, we would be committed to the conjoint reliability of well-confirmed theories of 'force', even when some of these theories are about military force and others about, say, mechanical force.

That a requirement of univocality must be part of the unity of science principle isn't surprising. What's of interest is the <u>epistemological</u> work which judgments of univocality must do.

From the fact that T' has been suitably experimentally tested (not using T'' as an auxiliary hypothesis, let us say), and the fact that T'' has been similarly tested (not using T' as an auxiliary hypothesis, let us say), and the fact that no term occurs ambiguously in $(T' \wedge T'')$ we are supposed to be justified in expecting the conjunct to be a reliable predictive instrument. What must univocality judgments be judgments <u>of</u> in order to play this sort of epistemological role?

It will hardly be surprising that an answer is readily available which extends the answer offered earlier when the issue of revision or extension of measurement procedures was at issue: Univocality for theoretical terms (contrary to the view of logical positivists) is sameness of reference. The principle of unity of science works because (assuming approximately true background theories, etc., as before) if T' and T'' are well confirmed then each is likely to be approximately true of the "theoretical" (and observational) entities which they are about. If none of their constituent terms occurs ambiguously (that is, with more than one referent) in $(T' \wedge T'')$ and if the respects in which T' and T'' approximate the truth are suitably compatible, then $(T' \wedge T'')$ will also be approximately true and that, in turn, is why its observational consequences will be approximately true.

The naturalistic account of the reliability of the principle of unity of science offered above, is--in a certain sense--incomplete without a corresponding account of the nature of reference. Moreover, the naturalistic and realistic accounts offered here and in section 2.1 impose constraints on theories of reference (at least for "theoretical terms"). Whatever the referential relation is between a theoretical term *t* and a theoretical kind (or property, or magnitude) *k*, it must be such that:

(1) It can be established and maintained by scientists who are engaging in the sort of practices which a realistic account of scientific methodology recommends, and

(2) judgments of univocality of the sorts which scientists ordinarily make must be reliable indicators of co-referentiality.

Let us turn from the question of reference for a moment and look more closely at the "micro-structure" of reliable-belief regulation in science as it relates to a particular theoretical kind (or magnitude, or property) *k*.

If the realistic and naturalistic account of scientific epistemology sketched earlier is correct, then the following are typical of the sorts of causal factors which make possible reliable social regulation of beliefs about *k*:

(i) Certain of the circumstances which are understood to be apt for the perception, detection or measurement of *k*(s) are, in fact, apt for the perception, detection or measurement of *k*(s).

(ii) Some of the circumstances which are taken to be indicative of certain features or properties of *k*(s) are, in fact, typically indicative of those features or properties of *k*(s).

(iii) Certain significant effects attributed to *k*(s) by experts are in fact produced by *k*(s).

(iv) Some of the most central of the accepted laws about *k*(s) are approximately true of *k*(s).

(v) There is some generally accepted description of k(s) which distinguishes it (them) from other kinds.

(vi) The socially recognized experts about k(s) are, in fact, members of an organized community whose beliefs about k(s) are reliably regulated.

All these factors, and others like them, serve to guarantee reliable k-belief regulation. That is, they tend to ensure that:

(vii) The sorts of considerations which rationally lead to modifications of, or additions to, existing k-theories are, typically and over time, indicative of respects in which those theories can be modified so as to provide more nearly accurate descriptions of k(s).

Consider now the following relations which might obtain between a term t and a kind k:

(i') Certain of the circumstances or procedures which are understood to be apt for the perception, detection or measurement of t(s) are, in fact, typically apt for the perception, detection or measurement of k(s). [Operationalism].

(Here, and in other entries on this list, I have abused the use-mention distinction. The reader will have no difficulty in providing (somewhat tedious) but correct reformulations of these points.)

(ii') Some of the circumstances which are taken to be indicative of certain features or properties of manifestations of t(s) are in fact typically indicative of those features or properties of manifestations of k(s). [Operationalism].

(iii') Certain significant effects attributed to t(s) by experts are in fact typically produced by k(s). [Putnam's example of theoretical term introduction by the citation of typical effects; see Putnam 1975].

(iv') Some of the most central laws involving the term t are approximately true if they are understood to be about k(s). [The "law-cluster" theory of meaning].

(v') There is some generally accepted putative definite description of t which is in fact true of k(s) and of no other kind, property or magnitude. ["Disguised definite description" theories of meaning].

(vi') The socially recognized t-experts form an organized community whose t-beliefs are so regulated that they tend to be true when they are understood to be about k(s). [Putnam's "division of linguistic labor"; see Putnam 1975].

(i')-(vi') are related in that, when many or all of them obtain to a significant extent with respect to a term t and a kind k, they will tend to bring it about that:

(vii') The sorts of considerations which rationally lead to modifications of, or additions to, existing theories involving the term _t_ are, typically and over time, indicative of respects in which those theories can be modified so as to provide more nearly accurate descriptions, when the term _t_ is understood as referring to _k_(s).

I have indicated after (i')-(vi'), above, theories of meaning or reference for theoretical terms which focus on relations between terms and kinds of the sort in question. It will not have escaped the reader that (i')-(vi') are nicely parallel to the micro-constituents of reliable belief regulation indicated in (i)-(vi). Indeed (i')-(vi') _re-state_ (i)-(vi) as constraints on the linguistic behavior of the scientific community if _t_ refers to _k_.

It will hardly be a surprise that I now propose that (i')-(vi') are typical micro-constituents of the relation of reference between a term _t_ and a kind _k_. Indeed, I suggest the sorts of relations between a term _t_ and a kind _k_, which would tempt us to say that _t_ refers to _k_, themselves form a "natural kind" just in virtue of the fact that they are the sorts of relations between the social uses of a term, _t_ and manifestations of a kind _k_ which--when enough of them are manifested--tend to bring it about that reliable _t_-belief regulation of the sort described in (vii') obtains.

I have elsewhere called relations like (i')-(vi') relations of "epistemic access" and I have suggested that it is just such relations which are the constituents of reference [Boyd 1979; see also Boyd forthcoming (b)].

In the context of the preceeding discussion of the naturalistic epistemology of science, the epistemic-access theory of reference can be put this way:

THE CONSTITUENTS OF RELIABLE BELIEF REGULATION ARE THE SAME AS THE CONSTITUENTS OF REFERENCE. KNOWLEDGE AND REFERENCE (BETTER: KNOWING AND REFERRING) HAVE THE SAME "MICRO-STRUCTURAL" COMPONENTS.

It is evident that the epistemic access account of reference satisfies the constraint indicated above that it explain how scientific practice establishes relations of reference of the sort desired. A moment's reflection will show that it also satisfies the constraint that it explain why our ordinary standards of univocality are indicative of co-referentiality. Each of the standard indicators of univocality suggested in the philosophical literature (sameness of "operational definition", sameness of "law-cluster", etc.) corresponds to an important component of epistemic access.

Let T'(_t_) and T''(_t_) be two theories containing the same lexicographic term _t_. The question of whether these two sorts of occurrences of _t_ are co-referential amounts to the question of whether--

in the practice of the relevant scientific communities--there is a single kind *k* such that this term *t* affords epistemic access to *k* in *both* the sort of research enterprise which gave rise to T' and the sort of research enterprise which gave rise to T''. If theoretical knowledge is possible at all, then it is no mystery that we can answer this sort of question reliably.

It is an intended consequence of the epistemic access account of reference that--under certain circumstances--relations of epistemic access can obtain between a term *t* and more than one kind and that such terms can come to be "disambiguated" as a result of the subsequent discovery of this fact. Field's notions of "partial denotation" and of "denotational refinement" [Field 1973, 1974] are thus special cases of the doctrine presented here. Moreover, it is a consequence of the present account that relations of epistemic access can obtain --and thus establish a reference-like relationship--even though it would be odd to say that full-blown reference obtains. Thus the present account can explain the grain of truth in such statements as: "When, in pre-scientific societies, people talk about various sorts of gods, it's really natural laws that they are talking about."

It remains to see that the phenomenon of reference displays the pattern of dialectical accommodation by successive approximation which is characteristic of scientific knowledge and scientific methodology. In the first place, the phenomena of partial denotation and subsequent denotational refinement illustrate the fact that the "tightness" of fit between language and the world increases as the dialectical process of theory refinement proceeds. The possibility of epistemologically useful referential relations weaker than full-blown reference, of the sort illustrated in the last paragraph, provides a further illustration of this phenomenon.

Secondly, there is a dialectical relationship between the reference-establishing epistemic relations for various different scientific terms which is partly obscured by the abbreviated formulation of the epistemic access account which I have offered here. I have suggested that relations like (i')-(vi'), when they obtain to a significant extent, tend to establish the sort of reliable belief regulation indicated in (vii'). What is in fact true is that to establish (vii') for any particular term *t* and kind *k*, relations like (i')-(vi') must obtain with respect to a large number of terms and kinds. [This is just to repeat the claim that successful induction in science depends upon a body of relevantly approximately correct background theories and beliefs.] Thus there are complex relations of mutual dependence between the reference-establishing relations for the whole range of theoretical terms in science (and in everyday life, for that matter).

Finally, consider the case in which at a particular time a term *t* refers to a kind *k* but the relevant scientific community makes certain as yet undiscovered systematic errors of classification, classifying as *k*s certain things which do not belong to *k*. It is reasonable to ask what makes it true that *t* then refers to *k* rather than to the nominal kind consisting of all those things which would _then_ be classified under _t_. After all, _at the time in question_, the connection between _t_ and this nominal kind might seem to be closer than the connection between _t_ and _k_. Why is it to _k_ that the use of the term _t_ then affords epistemic access?

The answer lies in the character of the mechanisms of reliable belief regulation themselves. It is because those mechanisms (which _are_ the mechanisms of epistemic access) depend upon, and establish, a dialectical process of accommodation by successive approximation between our _actual_ classificatory practices and the causal structure of the world, and because the reliability of these mechanisms depends crucially upon such a process of accommodation, that we are justified in thinking of those mechanisms as connecting _t_ to _k_ rather than to the nominal kind in question. The relation we describe when we say that a term refers to a kind just _is_ the relevant dialectical relation of accommodation.

3. Applications to Issues in the Philosophy of Science

In the following sections, I will describe ways in which the account of scientific knowledge sketched in the preceeding sections can be brought to bear on a number of issues in the philosophy of science. These discussions will be brief, and are intended only to indicate in broad outline how the distinctive features of that account are related to more specific philosophical issues.

3.1. The Refutation of Verificationism

The standard verificationist argument against realism goes (roughly) like this: Let T be any scientific theory which appears to describe unobservable phenomena. There will always be some other theory T', which has the same observational consequences as T, and which, if interpreted realistically, presents a different picture of unobservable phenomena. Since each theory would be equally well supported by any observational evidence, no evidence could determine which of these theories presents the correct account of unobservable phenomena. Therefore, scientific knowledge cannot extend beyond knowledge of observables.

There are two standard rebuttals to this sort of argument: (1) It fails to take account of the role of "auxiliary hypotheses" in theory confirmation, and thus it fails to recognize that T and T' might yield different observational consequences when (perhaps as yet undiscovered) auxiliary hypotheses are employed; (2) It depends upon a distinction between observable and unobservable phenomena; but no sharp distinction of that sort can be drawn. Neither of these rebuttals really constitutes an adequate reply to the epistemological thrust of the verificationist argument.

The first of these rebuttals is inadequate because the standard verificationist argument can be reformulated so that the theories in question are "total sciences"; that is, so that the theoretical commitments of the entire body of accepted scientific theories at a particular time are to be contrasted with those of a predictively equivalent alternative body of theories embodying a rival theoretical conception. Since "total sciences" contain all their own auxiliary hypotheses, the first rebuttal has no force. Such a total-science verificationism can even be extended to the consideration of rival theoretical traditions (Boyd 1973).

The second rebuttal fails because, in the first place, the significance of the verificationist argument does not depend upon a sharp observation-theory dichotomy. All that is required is a principled, if fuzzy, distinction which treats, say, atoms and "elementary" particles as unobservables. Moreover, there are at least two epistemologically plausible candidates for this principled distinction. One can either take as observable only those phenomena which are plainly observable to the unaided senses, or one can extend the notion of observability to those entities which can be detected when the senses are aided by devices whose relevant operation can be understood without reference to theories which themselves employ suspect terminology, like, magnifying glasses, and like optical telescopes and light microscopes when these instruments are employed in a theoretically unsophisticated manner (e.g., without spectroscopy equipment, polarizing filters, etc.). Of course, if one already knows that realism is sound, _then_ even this distinction will seem arbitrary, but it is certainly well-motivated with respect to the epistemological concerns which verificationism addresses, and its application in the verificationist argument is not rebutted by showing that if that argument is unsuccessful then the distinction is ill-motivated.

What is wrong with the verificationist argument, if the account of scientific epistemology defended here is correct, is that--under the considerations envisioned in the verificationist argument--the theories (or total sciences) T and T' are not necessarily equally well supported by scientific evidence. The evidential support for a theory --or the evidence against it--is not captured just by the confirmation or disconfirmation of the theory's observational predictions. Plausibility in the light of the theoretical tradition is also _evidential_; indeed considerations of theoretical plausibility of this sort are essential in assessing the strength of more "direct" experimental evidence. Furthermore, not just any theoretical tradition will do: the reliability of the experimental method depends upon its operations being governed by a relevantly approximately true theoretical tradition. If T is the actual current total science, and T' is an alternative total science which is profoundly implausible in the light of T, then there are _evidential_ reasons for preferring T to T'. The two total sciences are not equally well supported by available evidence.

The standard verificationist argument rests upon a highly plausible but mistaken interpretation of the (true) doctrine that factual knowledge is always grounded in observation. It is the naturalistic and contingent rebuttal to this interpretation which constitutes the deepest rebuttal to verificationism.[3]

3.2. The "Humean" Account of Causal Relations

The most durable empiricist doctrine in the philosophy of science must be the analysis of causal statements according to which the meaning of the claim that one event, e_1, caused another event e_2, is, roughly, that a statement asserting the subsequent occurrence of e_2 (or some appropriate statement about the probability of e_2's occurrence) can be deduced from natural laws, together with suitable descriptions of e_1's prior occurrence, and of relevant background conditions.

What makes this contemporary analysis Humean is that it follows Hume in rejecting the view that causal relations are to be understood as manifestations of natural necessity, or of the operation of causal powers or underlying mechanisms. The contemporary justification for a regularity analysis of causal relations is verificationist: what's wrong with understanding causation in terms of natural necessity or causal powers is that neither natural necessity nor causal powers are observable. Such a justification is also Humean: it captures the epistemological (but not the conceptual or psychological) reasons which Hume offered for the ancestor of the current doctrine.

The contemporary "Humean" analysis of causation bears a peculiar relation to its verificationist justification: it is quite often employed by philosophers who adopt a realist stance with respect to "theoretical entities"; indeed, it is not uncommon for the Humean analysis to be applied in cases where the events in question are *themselves unobservable*! The abandonment of the verificationist justification for the analysis seems not to have substantially reduced its acceptability.

Nevertheless, the Humean analysis is unacceptable from a realist perspective. If we have knowledge of unobservable entities and their unobservable properties--if we know, say, something about atoms, their properties and their sub-atomic constituents--then we *do* have knowledge of the underlying causal powers or mechanisms which manifest natural necessity. Such "secret powers" and "inner constitutions" of matter have always been the paradigm case of the sort of alleged causal reality against which the Humean account of causal relations has been directed. There is no longer any philosophical justification for the regularity analysis of causation.

It might seem otherwise, even to a realist. Consider the following puzzle: Suppose that realists are right and we know that matter is composed of various micro-constituents, with various unobservable properties and magnitudes and obeying certain (perhaps statistical) laws. These laws indicate how the various properties and magnitudes vary over time. They establish correlations between various sorts of microscopic physical states. But don't we need something like the Humean analysis of causal statements in order to justify our saying that these correlations are causal? Wouldn't we anyway need some alternative analysis of the meaning of causal statements in order to explain why we can understand scientific laws and theories of any sort as being causal, or as describing causal relations? If one replies that the relation between the relevant physical states is causal because, realistically understood, the laws in question describe the interactions between various constituents of matter and their respective properties and magnitudes, wouldn't that be circular, since the notion of interaction is itself a causal notion?

The answer is that what justifies our holding that the laws and relations in question are causal is--at least primarily--that paradigm cases of macroscopic causation are described by our theories as being composed of interactions between the micro-constituents of matter and their various properties. It is true that both the notion of interaction and the notion of composition (at least as it is employed here) are themselves causal notions. Thus the answer offered here would be circular if it were offered as a reductive definition of causation in non-causal terms. But, if causal relations and their constituents are real phenomena (as realism suggests), then there is no reason to believe that any such definition is possible. Indeed, it is doubtful that there are analytic definitions of central causal notions, even in terms of other causal notions. What we can reasonably expect is not an analysis of causation (in the philosopher's sense) but rather an assay of causation, an account of the sorts of causal factors there are in the world and of how they interact. That is the task of the various special sciences, of which philosophy is only one.

3.3. Theoretical Reduction and Theoretical Equivalence

Philosophers often need to ask what it is for two theories to be equivalent, to--in some sense--say the same things about the world. They must also ask what it is for one theory to be subsumable under another, to say only things about the world which the other theory already says. There is a standard answer in the positivist tradition to these questions: What it is for a theory T' to be subsumable under another theory T'' is for T' to be syntactically reducible to T'', i.e., for all the sentences in T' to be deducible from T'' when T'' is supplemented by a suitable set of definitions or "reduction sentences" relating the vocabularies of the two theories. Two theories are equivalent if each is syntactically reducible to the other.

Questions regarding relations of equivalence or subsumption between theories arise in a number of philosophical contexts. In some, the issue is ontological: what is at issue is, at least roughly, whether the entities countenanced by one theory would exist in a world in which the other is true. Questions of this sort arise in the philosophy of science in at least two quite different settings. In the first place, it has been a standard concern of philosophers of science to explicate materialist doctrines of, e.g., mental phenomena, or biology. It is plausible to hold that a significant component of such doctrines is the claim that the theories of certain special sciences are ontologically subsumable under the laws of fundamental physics.

Questions of ontological equivalence of theories also arise in certain discussions about limitations on possible theoretical knowledge. Philosophers sometimes inquire whether there might be theories which are in some respect or other ontologically inequivalent but which could not be distinguished by theoretical evidence. Now, the standard verificationist argument discussed in 3.1 addresses this issue, and finds that such situations are ubiquitous. But one need not adopt a verificationist perspective or a non-realistic conception of scientific evidence to raise such a question. It is entirely compatible with the sort of realism defended here that there should be some theoretical issues which, as a consequence of features of the actual laws of nature, could never be resolved by scientific investigation. (Moreover, to conclude that such an issue is thus unresolvable is not to adopt a "non-realistic" stance towards it, as some careless use of philosophical terminology seems to suggest.)

Nevertheless, there is a respect in which verificationist and anti-realist positions obscure the real issues in such discussions of ontological matters. The analysis of the relevant ontological notions of subsumption and equivalence in terms of syntactic reducibility is--if scientific realism is correct--a significant mistake. Consider the ontological question of the subsumption of a theory T' under a theory T''. What is at issue is whether or not the interactions of the entities, properties, events, magnitudes and causal powers described by T'' realize or constitute the entities, properties, events, magnitudes, and causal powers described in T', in the way in which, say, the interactions of the entities, magnitudes, etc. countenanced by physical theory constitute or manifest a chair, a table, or a thunderstorm. Now, it is clear that the notion of constitution or manifestation at issue is a causal notion; if ontological subsumption of this sort is analyzed according to the "Humean" account of causation, the result will be some version of the claim that T' is syntactically reducible to T''. Indeed, the "rational reconstruction" of materialist theories as doctrines about the syntactic reducibility of the special sciences to physics is the paradigm case of verificationist reconstruction of scientific doctrines.

But we have good reason to believe that the "Humean" analysis of causal relations is mistaken. Causal relations like constitution, composition or manifestation are real aspects of natural phenomena, and talk about them is not reducible to non-causal talk. In addition, there are more specific reasons to believe that syntactic reducibility is not an adequate analysis in these cases. In the first place, there are mathematical reasons for doubting that deductive subsumption under a law adequately captures the notion of causal determination even in cases where strict determinism obtains (Boyd 1972).

Moreover, there is good evidence to suggest that the syntactic analysis of ontological subsumption fails in just those cases which it was designed to account for. There is very strong evidence favoring the ontological claims of materialism in biology, and substantial evidence favoring such claims about mental phenomena. But there is at the same time evidence that neither biological nor psychological terms are physically definable in the way which syntactic reduction requires (Fodor 1974; Putnam 1967; Wimsatt 1976, 1979). Thus, the "Humean" reconstruction of these materialist doctrines transforms them from well-confirmed scientific hypotheses into doctrines which are almost certainly false. There is no reason to believe that the "Humean" reconstruction of theoretical equivalence in terms of mutual syntactic reducibility is any more satisfactory.

3.4. Realism, Paradigms, and Paradigm Change

In addition to straightforwardly verificationist arguments against realism, there are considerations involving the role of theoretical paradigms in science or the character of changes in paradigm which have seemed to some philosophers to mitigate against scientific realism. Some of these considerations reflect Kuhn's distinctly neo-Kantian conception of science (Kuhn 1970) whereas others are less neo-Kantian than just anti-realist. It is beyond the scope of this essay to survey these considerations in any detail (see Boyd 1979 for a fuller treatment of Kuhn's positions), but I will sketch the outline of a realist response to three of these anti-realist arguments:

(1) So paradigm-dependent is the methodology of normal science that, if it's supposed to represent a procedure for discovering facts about the world, then the world had better be paradigm-dependent as well.

(2) The change in world-view during major scientific revolutions is so great that both the meanings and the referents of theoretical terms change, so that the realist picture of science as producing successive approximations to the truth cannot be sustained.

(3) Even if a suitable account of referential continuity for terms in successive scientific theories is available, it will still turn out that the changes in world-view during scientific revolutions are so great that there will be no non-contrived notion of approximate truth which will permit us to describe earlier theories as approximately true and to sustain the picture of scientific progress by successive approximation as the realist conception of science requires.

(1) and (2) seem to be central features of Kuhn's anti-realist position. (2) is explicitly maintained from chapter IX on, in Kuhn 1970, and is most carefully articulated and defended in the well-known

discussion of the alleged deducibility of Newton's laws from Einstein's (Kuhn 1970, p. 101-102). (2) is my rendering of an important epistemological thread in Kuhn.

(3) represents the sort of anti-realist argument which might be advanced by someone concerned with the dramatic character of recent theoretical innovations in physics whether or not she adopts a neo-Kantian position in the philosophy of science. Thus, (3) might be defended on the following grounds: According to the classical atomic theory, matter is composed of discrete fundamental particles, which possess quite definite dynamical properties and which are quite unambiguously individuated. The quantum mechanical conception of matter presents a very different ontological picture. Not only do the classical dynamical variables appear to lack simultaneous sharp values, but it is hard to see how the quantum mechanical formalism, as it is usually interpreted, reflects a theory of discrete particles at all. If, as realists insist, we are to take the quantum mechanical picture as an approximately faithful representation of reality, then it is hard to see how, as realists also insist, the previous classical theory was also an approximately true description of the same unobservable reality. The ontological picture presented by the two successive theories is just too different. Worse yet, what it is for a theory to be approximately true is for it to say approximately true things about the entities which it countenances; but on the quantum mechanical view, there are no such fundamental particles as the classical theory accepts. Thus, unless one can defend an appropriate non-standard interpretation of the quantum mechanical formalism, realism seems to be refuted.

The account of scientific realism presented in this essay permits us to see insights in each of these anti-realist arguments, while simultaneously defending the realist conception of scientific progress. With respect to (1), the realist account offered here concurs that a special explanation is required for the reliability of the profoundly paradigm-dependent methodology which characterizes mature sciences. It sees the explanation as lying in the contingent emergence of a relevantly approximately true body of background theories rather than in the paradigm-dependence of the world which scientists study.

Kuhn's arguments for (2) consist largely in maintaining a version of the law-cluster account of the reference and meaning of theoretical terms according to which certain central features of the relevant paradigm are part of the meaning of theoretical terms and serve to fix their referents. Kuhn holds that changes in these paradigmatic features indicate a change in reference. The account of reference offered here concurs that the fundamental features of the paradigm are the relevant reference-fixing mechanisms, since these are just the belief-regulating mechanisms appealed to in the epistemic access account of reference. But because the naturalistic account of reference offered here sees these mechanisms as establishing a causal connection between theoretical terms and their referents, rather than as embodying conceptual truths, it does not follow from that account that the paradigm must embody exactly correct descriptions of the relevant theoretical entities. Therefore, a change in paradigm need not indicate a change in the referents of the relevant theoretical terms.

With respect to (3), the dialectical and naturalistic account of the progress of scientific knowledge offered in the present essay permits us to say in just what respects the scientific realist should, _prima facie_, expect the process of successive approximation to be reflected in the transition from one paradigm to its successor in a mature science: (a) The terms of the preceeding theory should, for the most part, correspond systematically (though sometimes ambiguously) to real features or aspects of reality, and thus to features or aspects of the phenomena which the successor theory countenances; (b) the relation of correspondence should be constituted by relations of epistemic access of the sort which the realist sees as constituents of the mechanisms of reliable belief regulation; (c) the existence of this correspondence should explain the (instrumental and theoretical) success of scientific methodology under the earlier paradigm; (d) the theoretical knowledge which is reflected in the preceding paradigm should help to explain both how its successor was discovered and how it was confirmed: that is, the theoretical knowledge embodied in the earlier theory should help to explain the reliability of the belief-regulating mechanisms which governed the transition between the paradigms; (e) finally, (a)-(d), which are themselves empirical claims, should be features of the picture of reality afforded by the successor paradigm (together with the relevant historical and philosophical theories).

No novel or non-standard interpretation of any recent scientific revolutions is necessary in order to defend the sort of theoretical continuity represented by (a)-(e). The "overlap" between successive paradigms in the 20th-century, for example, is of exactly the sort required for their defense. Nevertheless, there is an important insight behind the third objection to realism which we have been considering. It appears to be the case that--as a matter of fact--we are not (or, perhaps, not yet) very good at identifying the _fundamental_ features of nature, at least in physics. Features of reality which are treated by one theory as fundamental are often treated by its successors as aspects of much more complex phenomena. Terms which seem to correspond to quite definite properties may turn out to correspond to aspects or features of some more complex sort of physical state. Similarly, terms which seem to correspond to discrete and clearly individuated entities may instead refer to aspects of a more complicated reality. Insofar as we think of "fundamental physics" as seeking to describe the ultimate and fundamental features of physical reality, then its efforts appear to have been, thus far at least, unsuccessful. Whatever the explanation for that failure, however, there is no reason from the recent history of physics to deny that the methods of physicists are reliable with respect to the discovery of _approximate_ truths about real unobservable features of matter.

Whenever, in this century, scientists and philosophers have examined features of existing physical theories which seem scientifically or philosophically puzzling or which suggest that our deepest theories may be in some respect inadequate or incomplete as descriptions of underlying physical reality, some of them have been tempted to see the apparent inadequacy of our current theories as dictating an abandonment of scientific realism. The abandonment of realism is a great theoretical-puzzle elimination strategy: if theories are never supposed to describe underlying reality, then it's no puzzle that the current ones don't seem fully adequate to the task. Indeed. The abandonment of realism is the abandonment of theoretical inquiry. Even without the epistemological picture presented here--which makes instrumental knowledge parasitic on theoretical knowledge--the cure seems worse than the disease. If the epistemological conception of the present essay is sound, then the abandonment of theoretical inquiry would entail the abandonment of the search for instrumental knowledge as well. Nothing in the recent history of science would justify the abandonment of either sort of inquiry.[4]

Notes

[1]The views in this paper were developed over the last decade. I have, therefore, had the opportunity to profit from discussions with a great many people, some of whom will not even remember our conversations. I want especially to thank William Alston, Ned Block, George Boolos, Sylvain Bromberger, Richard Cartwright, Norman Daniels, Hartry Field, Alan Gilbert, Carl Ginet, Alvin Goldman, Alex Goldstein, Kristin Guyot, Harold Hodes, Paul Horwich, Hilary Kornblith, Barbara Koslowski, Thomas Kuhn, Richard Miller, Henry Newell, Andrew Ortony, Mark Pastin, William Provine, Hilary Putnam, Israel Scheffler, Sydney Shoemaker, George Smith, Robert Stalnaker, Howard Stein, Nicholas Sturgeon, Robert Weingard, William Wimsatt and David Zaret.

Various parts of this paper were presented in graduate seminars at Cornell and at M.I.T. and in colloquia at Case-Western Reserve University, Princeton, The Rockefeller University, Rutgers, The State University of New York at Oswego, Syracuse University, Tufts, The University of Minnesota, The University of Chicago, The University of Illinois at Chicago Circle, and the University of Illinois at Urbana-Champaign. I thank the audiences at these presentations for their helpful comments and criticisms.

[2]The considerations rehearsed here provide a rebuttal to a natural objection to the argument for realism offered in 2.1. The natural rebuttal goes like this: The methodological principles which characterize mature science are as paradigm dependent as the realistic account of scientific epistemology suggests, and their reliability seems to depend upon their paradigm dependent features. But, all we can conclude is the following inductive generalization: Scientific research conducted according to such a method (i.e., scientific research conducted as a realist would recommend) is, and will be, instrumentally reliable. What is unwarranted is the "metaphysical" realist

explanation of this generalization.

The realist replies by raising the issue of projectability for the characterization of scientific methodology which the non-realist accepts: There are infinitely many possible "methodologies" which would have recommended the finitely many methodological judgments which have characterized the practice of science thus far, but which (pairwise) differ about future methodological recommendations. What reason have we to believe that the sort of description of methodology employed in the non-realist's inductive generalization about induction is the appropriate sort of description? What reason have we to believe that this sort of description captures the reliability-making features of past scientific practice? It's true that other characterizations of scientific method to date would be arbitrary, or silly, but what would their arbitrariness or silliness consist in?

The answer, I suggest, is that our initial confidence in the description in question already rests upon the common-sense, pre-philosophical, realistic understanding of the principles involved, and of the reasons why they are justified. The second-order inductive generalization about inductive methods offered by the non-realist rests upon a judgment about projectability of descriptions of past methodological practices which itself reflects a tacit acceptance of the sort of first-order theoretical inductions involved in 2.1. When this tacit support is withdrawn, the burden of proof rests upon the non-realist to justify her proposed identification of the reliability-making features of previous scientific practice.

It is worth remarking that what is illustrated here is a quite general feature of anti-"metaphysical" empiricism. Empiricists deploy skeptical arguments which, if consistently developed, would cast doubt upon all general knowledge whatsoever. The exceptions they make for general knowledge of observable phenomena are ultimately unjustifiable within empiricist epistemology (see section 3.1).

[3]This rebuttal to verificationism reflects an important difference between the strategy for defending realism employed in 2.1, and a more common strategy for defending realism which philosophers of science often employ. According to this second strategy, realism is defended by arguing, for particular scientific theories, that the best explanation for their instrumental reliability is the approximate truth of the laws they contain. This latter strategy provides us with good reasons to believe that realism is true, and that there is, therefore, something wrong with the verificationist epistemological arguments against the possibility of theoretical knowledge. It does not, however, tell us what is wrong with verificationist epistemology.

The argumentative strategy of 2.1, by contrast, provides an epistemological defense of realism, permits us to articulate a distinctly realistic understanding of the epistemology of science, and makes it possible to say just what is wrong with the central epistemological argument of verificationism.

[4]It should be noted that neither the argument of 2.1 nor the discussion of approximate truth in 3.4 (nor both together) entail that the methods of actual science would lead to exactly true theories as an "asymptotic limit" if science were pursued long enough. Nothing in the arguments for realism presented here is meant to preclude the possibility that there are true theories which we could never discover, or issues we would never get exactly right. Over successive approximations to the truth could be "bounded away" from the exact truth for any number of reasons--intellectual limitations, causal limitations on access to data, for example, or some systematic and irremediable defect in our conceptual or theoretical framework. What I do insist is that--contrary to what logical positivists have claimed--we have lots of approximate knowledge about unobservable phenomena and our ability to improve our instrumental knowledge is largely parasitic upon our ability to improve our theoretical knowledge.

The account of science offered here is, therefore, not some sort of realist's version of the pragmatic definition of truth as the limiting case of rational scientific investigation. In the language of Putnam 1978 (see Part Four) exact truth is "radically non-epistemic", even though reference--from which truth can be defined a la Tarski--is itself an epistemic notion. Truth (or, rather, exact truth) is disconnected from our rational methods in a way in which reference and approximate truth are not. (I am grateful to William Wimsatt and Howard Stein for discussions which led to the inclusion of this note.)

References

Boyd, R. (1972). "Determinism, Laws and Predictability in Principle." Philosophy of Science 39: 431-450.

-------. (1973). "Realism, Underdetermination and a Causal Theory of Evidence." Nous 7: 1-12.

-------. (1979). "Metaphor and Theory Change." In Metaphor and Thought. Edited by Andrew Ortony. Cambridge: Cambridge University Press. Pages 356-408.

-------. (1980). "Materialism Without Reductionism: What Physicalism Does Not Entail." In Readings in Philosophy of Psychology, Volume 1. Edited by Ned Block. Cambridge: Harvard University Press. Pages 67-106.

-------. (forthcoming (a)). The Physical Basis of Mind. Cambridge: Harvard University Press.

-------. (forthcoming (b)). Realism and Scientific Epistemology. Cambridge: Cambridge University Press.

Feigl, H. (1956). "Some Major Issues and Developments in the Philosophy of Science of Logical Empiricism." In The Foundations of Science and the Concepts of Psychology and Psychoanalysis. (Minnesota Studies in the Philosophy of Science, Volume 1.) Edited by H. Feigl and M. Scriven. Minneapolis: University of Minnesota Press. Pages 3-37.

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

--------. (1974). "Quine and the Correspondence Theory." Philosophical Review 83: 200-228.

Fodor, J.A. (1974). "Special Sciences, or the Disunity of Science as a Working Hypothesis." Synthese 28: 97-115.

Goldman, A. (1967). "A Causal Theory of Knowing." Journal of Philosophy 64: 357-372.

----------. (1976). "Discrimination and Perceptual Knowledge." Journal of Philosophy 73: 771-791.

Hempel, C. (1965). Aspects of Scientific Explanation. New York: The Free Press.

Hume, D. (1739). A Treatise of Human Nature. London: John Noone. (As reprinted (ed.) L.A. Selby-Bigge. Oxford: Oxford University Press, 1888.)

-------. (1748). Philosophical Essays Concerning Human Understanding.

London: Printed for A. Millar. (As reprinted as An Inquiry Concerning Human Understanding. (ed.) Charles W. Hendel. New York: Bobbs-Merrill, 1955.)

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

Locke, J. (1690). An Essay Concerning Human Understanding. London: Thomas Basett. (As reprinted from the 5th edition (1706). (ed.) John W. Yolton. New York: Dutton, 1961.)

MacCorquodale, K. and Meehl, P. (1948). "On a Distinction Between Hypothetical Constructs and Intervening Variables." Psychological Review 55: 95-107.

Nagel, E. (1961). The Structure of Science. New York: Harcourt Brace.

Pastin, M. (1978). Review of Armstrong, D.M. Belief, Truth and Knowledge. Metaphilosophy 9: 150-162.

Putnam, H. (1967). "The Mental Life of Some Machines." In Intentionality, Minds and Perception. Edited by H. Castañeda. Detroit: Wayne State University Press. Pages 177-200.

---------. (1975). "The Meaning of 'Meaning'." In Language, Mind and Knowledge. (Minnesota Studies in Philosophy of Science, Volume VII). Edited by K. Gunderson. Minneapolis: University of Minnesota Press.

---------. (1978). Meaning and The Moral Sciences. London: Routledge and Kegan Paul.

Quine, W.V. (1970). "Natural Kinds." In Essays in Honor of Carl G. Hempel. Edited by Nicholas Rescher. Dordrecht: D. Reidel. Pages 5-23. (As reprinted in Ontological Relativity and Other Essays. New York: Columbia University Press, 1969. Pages 114-138.)

Smart, J.J.C. (1963). Philosophy and Scientific Realism. London: Routledge and Kegan Paul.

Wimsatt, W.C. (1976). "Reductionism, Levels of Organization, and the Mind-Body Problem." In Consciousness and the Brain. Edited by G.G. Globus, G. Maxwell and I. Savodnik. New York: Plenum. Pages 205-267.

------------. (1979). "Reduction and Reductionism". In Current Research in Philosophy of Science. Edited by P.D. Asquith and H. Kyburg, Jr. East Lansing, Michigan: Philosophy of Science Association. Pages 352-377.