Lex Orandi est Lex Credendi

Richard N. Boyd

> But whilst we are destitute of senses acute enough to discover the minute particles of bodies, and to give us ideas of their mechanical affections, 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 natural bodies: and our reason carries us herein very little beyond particular matters of fact. Locke, *Essay*.

Van Fraassen's *The Scientific Image* (1980) represents a sophisticated and brilliant defense of empiricist philosophy of science. Among the many interesting features of this work are van Fraassen's rebuttals to a class of arguments for scientific realism which other philosophers and I have offered over the past decade or so and which may be characterized as efforts to defend scientific realism as an empirical hypothesis which is justified because it provides the best scientific explanation for various facts about the ways in which scientific methods are epistemically successful. These arguments capture an important source of the plausibility of scientific realism. For that reason, it is important to assess the extent to which these arguments are successful against sophisticated empiricism. Van Fraassen's rebuttals represent just the sort of sophisticated empiricist criticisms against which these arguments should be tested. The aim of the present paper is to show that these rebuttals are ultimately unsuccessful and that the arguments for realism can be sustained.

In section 1 of this paper, I present in some detail empirical arguments for realism of the sort in question. In section 2, I discuss criticisms of

A version of this paper was presented at the philosophy colloquium at The Johns Hopkins University. I thank the audience, especially Peter Achinstein, Gary Hatfield, and David Zaret, for their helpful comments and criticisms. I have also profited greatly from discussions with John Bennett, Philip Gasper, Kristin Guyot, Harold Hodes, Christopher Hughes, Norman Kretzmann, Scott MacDonald, Robert Rynasiewicz, Sydney Shoemaker, Alan Sidelle, and Allen Wood. Rega Wood provided very helpful bibliographical advice.

those arguments which are presented by van Fraassen or are suggested by points which he makes. In section 3, I draw some lessons about philosophical and scientific methodology.

I. Realism and the Theory-Dependence of Experimental Design

In several papers (Boyd 1983, 1981, and esp. 1973), I have argued that scientific realism provides the only scientifically reasonable explanation for the reliability of certain important features of scientific methodology which are crucial in experimental design and in the assessment of experimental evidence. Roughly speaking, these are the features of scientific methodology relevant to the assessment of the "degree of confirmation" of a proposed theory, given a body of observational evidence (if we choose to employ the standard empiricist terminology). In the present section, I want to expand upon this claim and to offer arguments for it in somewhat greater detail.

To begin with, it is important to understand what sort of reliability of methodology is to be explained. If scientific realism is true, then the methodological practices of science provide a reliable guide to approximate truth about theoretical matters and, no doubt, only scientific realism could provide a satisfactory explanation for this fact. But it would be question-begging to suggest that this provides any good reason to accept scientific realism; after all, only realists believe that the methodology of science is reliable in this sense, anyway. What I propose to do is to take advantage of the fact that antirealists in the philosophy of science are typically selective in their skepticism and to define the reliability of the methods of science in such a way that no questions are begged against the position of the typical antirealist. Call a theory instrumentally reliable if it makes approximately true predictions about observable phenomena. Call a methodology instrumentally reliable if it is a reliable guide to the acceptance of theories which are themselves instrumentally reliable. For the antirealist against whom my arguments are directed, it is uncontroversial that the actual methods of science are instrumentally reliable in this sense, although it may of course be a matter for philosophical dispute just which features of actual scientific practice explain this reliability. The arguments I am discussing here are directed against only the selectively skeptical antirealist; I have nothing to say to "the Skeptic."

Let us suppose that some scientific theory T has been proposed and that a body E of experimental results has been obtained which is consonant with the predictions of T. Imagine that T is an ordinary medium-sized theory of the sort which scientists routinely confirm or disconfirm in the course of what Kuhn calls "normal science." (Scientific realism must have

something to say about the acceptance of large-scale paradigm-fixing theories as well, but it is a matter of controversy whether there is a reliable methodology for such cases, and so I want here to examine the more commonplace instances of theory testing. For a realist treatment of the other cases, see Boyd 1979, 1981, 1983.) Questions regarding the extent to which T is confirmed by the evidence E may fruitfully divided into three categories:

1. *The question of "projectability."* One of the things which Goodman (1973) has taught us is that something important about inductive inference can be learned by examining the *unrefuted* inductive generalizations which no one ought to accept. Goodman formulates the issue in terms of the projectability of predicates in simple inductive generalizations, but it is clear that the issue he raises is more general. We can think of any sufficiently general theory as representing the proposal to consider as projectable certain possible patterns in observable data, *viz.*, those patterns which the theory predicts. In general, the methodological acceptability of such a proposal will not depend soley upon the projectability of the individual predicates contained in the theory in question considered in isolation, but also on the structure of the theory itself (see Boyd 1981, sec. 2.3; Boyd 1983, sec. 8, esp. p. 41). T will receive significant evidential support from E only if T represents a projectable pattern in possible observational data.

2. *The question of experimental controls and experimental artifacts.* Suppose that T represents a projectable pattern in possible observational data, and suppose further that the data in E represent apparently confirming evidence for T, whatever this latter constraint might come to. It will still be methodologically inappropriate to accept T as well confirmed unless there is reason to believe that the experiments involved in the production of these data were well designed. There must have been experimental controls for the influence of factors irrelevant to the assessment of T; in particular, the data which appear to confirm T must not be artifacts of the design of the experiments in question rather than genuine tests of the empirical adequacy of T. The analogous constraint applies, of course, to the case in which T is apparently disconfirmed by E.

3. *The question of "sampling."* Suppose that T represents a projectable pattern and that the experiments whose results are reflected in E are individually well designed. If we now ask how well, or to what extent, T is confirmed by E, we face head on the methodological analogue of the pure epistemologist's problem of induction. T will typically have infinitely many different observational consequences, and the problem of assessing the extent to which E confirms T comes down to the question of which

(typically relatively small) finite subsets of those consequences are such that their confirmation bestows significant confirmation on all the rest. (For the realist, of course, the problem is broader; one needs to know which such subsets bestow significant confirmation on the theory taken literally as a description of [partly] unobservable reality. Here, as in the case of the definition of reliability of methodology, I frame the issue in a way which does not beg the question against the antirealist.)

We may, I think, frame this question in a revealing way. The question is whether the consequences of T which have been tested are—in an epistemically appropriate sense—*a representative sample* of all the observational consequences of T. We cannot have checked out all of the (epistemically) possible ways in which T could "go wrong" with respect to observational prediction; there are, after all, infinitely many such ways. What we want to know is whether the experimental studies in question involve a representative sample of those ways, so that, if T hasn't gone wrong where we've tested it, then we can be justified in believing that it isn't going to go (very far) wrong at all. (Actually, this description is somewhat idealized; in the actual history of science, well-established theories have often turned out to be very wrong indeed in some of their empirical predictions. What is important is that we—rightly—expect experimental confirmation of a theory to warrant our belief that it will prove instrumentally reliable in a wide range of applications whose limits we cannot set in advance. The problem of identifying a relevantly representative sample of the observational predictions of a proposed theory is hardly rendered easier by this complication.)

The theory-dependence of the answers to these questions

The ways in which scientists answer these fundamental questions regarding the assessment of experimental evidence are quite profoundly dependent upon their prior theoretical commitments. That this is generally true of scientific methodology is now uncontroversial; Kuhn, Quine, both H. Putnams, Goodman, Glymour, and van Fraassen have all emphasized this point without, of course, all drawing realist conclusions. It will be important for our purposes to examine in some detail the ways in which theoretical considerations are involved in answering the three questions about experimental evidence which we have just identified.

1. *Projectability*. Kuhn (1970) correctly insists that in mature sciences the basic form of solutions to particular research problems is tightly circumscribed by the theoretical and research tradition (the "paradigm"), and van Fraassen (1980) agrees that the acceptance of particular theories involves the scientist "in a certain sort of research programme" (p. 12). The proposed theory T whose degree of confirmation by E is to be estimated will not be a serious candidate for comfirmation at all unless it arises as a proposed solution to some problem: the extension of an existing theory to some new area of application, perhaps, or the explanation of some particular phenomena or observations. The theoretical tradition very sharply constrains such proposals; a proposed solution is unacceptable unless it is *theoretically* plausible in the light of existing theories, unless it is one of the solutions suggested by the existing "paradigm." Only those patterns in observable data are considered projectable which correspond to theoretically plausible theoretical proposals. Two facts about the theory-dependence of such projectability judgments are important for our concerns.

In the first place, such judgments sharply limit the generalizations *about observables* which we take to be confirmable. Suppose, as is typically the case, that T is put forward to account for some particular (finite) set of observational data. Of course, there will be infinitely many possible theories which would accommodate those data. Even if we take two such theories to be equivalent if they are empirically equivalent (or, better, if their respective integrations into the existing theoretical tradition would be empirically equivalent), there will remain infinitely many equivalence classes, each representing one possible observational generalization from the initial data. The effect of our theory-dependent judgments of projectability is to restrict our attention to a quite small finite number of these possible generalizations. Only the generalizations in this small set are potentially confirmable by observations, given the prevailing standards for the assessment of scientific evidence.

Secondly, the projectability judgments in question are genuinely *theory*-dependent. The judgments of theoretical plausibility which these projectability judgments reflect depend upon the *theoretical* structure both of the proposed solutions and of the received theoretical tradition. Proposed problem solutions are plausible, for instance, when the unobservable mechanisms they postulate are relevantly similar to the mechanisms postulated in the received theoretical tradition, where the relevant respects of similarity are likewise dependent on the theoretical structures postulated in the tradition. If the received body of theories were replaced by some quite different but empirically equivalent body of theories, then judgments of theoretical plausibility would pick out quite different problem solutions as acceptable and thus typically identify quite different patterns as projectable. As Kuhn insists, the ontology of the received "paradigm" is crucial in determining the range of acceptable problem solutions (and thus the range of projectable patterns in data).

2. *Experimental artifacts*. Suppose that T is theoretically plausible and thus represents a projectable pattern in observable data, and suppose that the experimental results in E appear to support (or refute) T. If these results are really to be evidentially relevant, then there must be reason to think

that the results favorable (or unfavorable) to T were not the result of features of the experimental situation which are irrelevant to the assessment of T. Of course, it is impossible to control for all epistemically possible experimental artifacts (of which there is an infinite number). Instead, we rely upon established theory to indicate the conditions under which the presence of experimental artifacts is to be suspected and the sorts of experimental controls which will permit us to avoid or discount for their effects. This is, I think, uncontroversial. It is also uncontroversial—although it is not much stressed in the literature—that our theory-dependent judgments in this area cut down the number of epistemically possible artifactual effects we actually control for from infinitely many to rather few.

What may be more controversial is whether or not these judgments are theory-dependent in the broader sense that they depend on the *theoretical* structure of the relevant background theories rather than just on their observational consequences. It might seem that they do not. Consider, for example, the commonplace that one must, in experiments involving electrical phenomena, control for the 60Hz hum induced by the alternating current in ordinary electrical wiring. Of course, the background theories which draw our attention to this sort of possible artifact have a complex theoretical structure, postulating electrons and electrical and magnetic fields and so forth. But, in order to know that we must shield various pieces of apparatus, all we need to know is that unless we do there will appear a certain sort of signal in our recording equipment superimposed on whatever signal comes from the preparation we are studying. If this sort of situation always obtains in cases of controlling for experimental artifacts, then it would appear that the methodological judgments which govern such controls do not depend on the theoretical structure of the relevant background theories.

Even if this were the case, there would, of course, be a significant way in which the identification of necessary experimental controls depends on the *theoretical* structure of the theories in the relevant theoretical tradition: the judgments of "projectability" which governed the acceptance of the generalizations about observables reflected in the currently accepted theories would have themselves depended on the *theoretical* structure of the earlier stages in the theoretical tradition. More importantly, it is by no means the case that the identification of relevant possible experimental artifacts depends solely on the observational consequences of the relevant background theories. This is so for two related reasons. In the first place, sound methodology often requires that we control for possible experimental artifacts whose effects are not by any means *predicted* by the received body of the theories but whose interference with the intended function of the experimental apparatus is *suggested* by those theories. Whatever

may be the ultimate "rational reconstruction" of our practice, it is true that, in the typical case, the way in which the possible artifactual effects are suggested is that there are (typically unobservable) mechanisms postulated by the received theories about which it is *theoretically* plausible that either these mechanisms or mechanisms similar to them will produce the artifactual effects in question. Thus, we identify relevant possible experimental artifacts by something like "inductive" inference from theoretical premises, and the sorts of possible artifacts which we thereby identify depend dramatically on the theoretical structure of the theories which are the premises of these inferences.

We may see the same sort of theoretical-structure-dependent inferences in another methodologically important strategy for the identification of relevant possible experimental artifacts. Good methodology often requires that we control in one experimental situation E for some possible artifact A because we have already encountered similar artifacts A' in similar experimental situations E'. In the typical case, the possible artifact will be described in partly theoretical language, and the relevant respects of similarity (between E and E' and between A and A') will be determined by theoretical considerations—by considerations about the structure and effects of the unobservable mechanisms which the received theories postulate as operating in the relevant natural systems. In this case, too, whatever the ultimate reconstruction might be, the theoretical structure of the accepted theories, and not just their observational consequences, plays a crucial role in the identification of the relevant possible artifacts.

Two important points of similarity thus emerge about the way in which sound scientific methodology controls for the possibility of experimental artifacts and the way in which the problem of projectability is solved. In the first place, while there are infinitely many epistemically possible experimental artifacts which might affect any given experiment, scientific attention is paid to only a small finite number. In this regard, the identification of relevant possible experimental artifacts resembles the assessment of projectability: from an infinity of epistemic possibilities, the scientific method identifies a small finite number as methodologically relevant. The identification of relevant possible experimental artifacts resembles the solution to the problem of projectability in another crucial way: in each case the relevant methodology depends on the theoretical structure of the currently accepted scientific theories; were those theories replaced by others which are empirically equivalent but theoretically divergent, quite different methodological practices would be identified as appropriate. In both cases, scientists behave as though their methodology were determined by inductive inferences from the theoretical principles embodied in the received theoretical tradition.

3. *Sampling*. The pattern discernible in our examination of the ways in which scientific methodology handles the issues of projectability and experimental artifacts is even more striking in the case of the solution to the problem of "sampling." It is a fair statement of the most basic methodological principle governing the assessment of experimental evidence that a proposed theory *T* should be tested under conditions representative of those in which it is most reasonable to think that the theory will fail, if it's going to fail at all. The identification of these conditions rests upon *theoretical* criticism of that theory. The proposed theory *T* will, typically, postulate various mechanisms, entities, processes, etc., as factors in the phenomena to which it applies. Theoretical criticism involves the identification of alternative conceptions of the mechanisms, processes, etc., involved which are theoretically plausible—that is, which are suggested by the sorts of mechanisms, entities, etc., which are postulated by the received body of theories. These theoretically plausible alternatives to *T* will suggest circumstances in which the observational predictions of *T* might be expected to be wrong. It is under (representative instances of) these circumstances that *T* must be tested if it is to be well confirmed. This is the central methodological princple of experimental design.

Plainly, the methodological solution to the problem of sampling is theory-dependent. Moreover, whatever the "rational reconstruction" of this methodology might be, scientists do not in practice distinguish sharply between unobservable mechanisms, processes, entities, etc., and observable ones in identifying ways in which a proposed theory might reasonably be expected to fail. Indeed, inferences which look for all the world like inductive inferences from accepted premises about unobservables to conclusions about unobservables play an absolutely crucial role in the sort of theoretical criticism we are discussing. Thus, in the present case, as in the case of the methodological solutions to the problems of projectability and of experimental artifacts, the ways in which scientific methodology is theory-dependent are such that, if the existing body of theories were replaced by an empirically equivalent but theoretically divergent body of theories, our methodological judgments *regarding the "degree of confirmation" of generalizations about observables* would be profoundly different.

Projectability and induction about unobservables

The ways in which the features of scientific methodology just discussed depend on the theoretical structure of the received body of background theories may be seen more clearly if we consider a standard way in which philosophers in the tradition of logical positivism have treated the feature of theory testing which I have presented under the heading "Sampling." It has been widely recognized that at any given time in the history of science, and for any given problem or issue, there are typically only a very few theories "in the field" and contending for acceptance. The practice of testing a proposed theory against its most plausible rivals might, in this context, be seen as simply an application of the same pragmatic principle which dictates that, if there are only a very few brands of band saw available, one should evaluate each before making a purchase. This sort of description gives the appearance of reducing the methodological principle we have been discussing to a *merely* pragmatic level, denying it any special epistemic relevance.

Of course, such an interpretation *would not* deprive the practices we have been discussing either of their theory-dependence or of their epistemic importance. So long as the relevant rival theories are identified in the theory-dependent way described in the section on projectability, then the practice would be as theory-dependent as one could wish. Moreover, if testing proposed theories against instrumental rivals in the way suggested represented *the* methological solution to the problem of "sampling," then it has epistemic importance however much it may also have a purely pragmatic justification. What is most interesting in this context, however, is that the actual methodological practice of scientists departs in a revealing way from that suggested by the pragmatic account. In order for the pragmatic picture to have any plausibility, we must think of a proposed theory as competing against other possible predictive instruments roughly as powerful as itself. The rival "theories" against which it must be tested must be theories in the sense of fairly well developed systems with some significant predictive power. One does, after all, test band saws against other band saws.

While it is true that sound scientific methodology does require that a proposed theory should be tested against similarly well articulated rivals which are approximately equally theoretically plausible (roughly, that's what being a rival *theory* means, beyond having been invented in the first place), what is striking about the methodological practices which constitute the solution to the problem of sampling is that they may also require that a proposed theory be tested against a mere hunch, which has no deductive predictive consequences whatsoever. Suppose that a proposed theory *T* postulates a particular sort of unobservable mechanism as operating in the systems to which *T* applies, and imagine that *T* is sufficiently well worked out that (using well-established auxiliary hypotheses) it is possible to obtain experimentally testable predictions from *T*. Suppose, also, that theoretical criticism of *T* identifies alternative possible mechanisms, plausible in the light of received theories. Under these circumstances, it becomes necessary to try to pit *T*'s conception of the matter against the alternative in some sort of experimental situation *even if* the alternative account is not nearly so thoroughly worked out as *T* and even if, for that reason, it yields (together with relevant auxiliary hypotheses) *no* definite predictions about observables at all. Under these circumstances, what sound methodology dictates is the identification of experimental circumstances under which the sorts of observations which it is *theoretically plausible* to expect given the alternative conception are different from those which one would expect given *T*'s account of the relevant mechanisms.

By way of example, suppose that *T* provides an account of the reaction mechanisms for some biochemical process and that *T* is worked out in sufficient detail that it has (together with well-confirmed auxiliary hypotheses) significant deductive observational consequences. Suppose that a rival conception of the relevant reaction mechanisms is suggested by theoretically plausible considerations but that this rival conception is insufficiently well developed to have specific testable deductive consequences. It might nevertheless be possible to test *T* against the rival conception. Suppose that, in the case of better-studied systems, those systems to which mechanisms like those proposed by *T* and ascribed by the received theories are much more sensitive to some particular class of chemical agent than those to which mechanisms like those proposed in the alternative are ascribed (note here that the relevant respects of likeness will be determined by the content of theoretical descriptions of the systems in question, and the theoretical content of the relevant background theories, *and* that the class of chemical agents in question may similarly be theoretically defined). Under such circumstances, sound methodology will dictate subjecting the biochemical systems to which *T* applies to chemical agents in the relevant class; data indicating considerable sensitivity to such agents will be especially important for the confirmation of *T* precisely because they will constitute a test of *T* against the theoretically plausible rival conception of the relevant reaction mechanisms. Note that in the present case the rival conception need not have any *deductive* observational consequences regarding the experimental situations in question. Instead, reasoning by analogy *at the theoretical level* makes it *theoretically plausible* to expect low sensitivity if the rival conception is true. *T* is tested against a theoretically plausible hunch about how it might go wrong.

Indeed, the role of considerations of theoretical plausibility in theory testing can go even deeper; a proposed theory may be pitted against a theoretically plausible rival in a particular experimental setting even though neither the rival *nor* the proposed theory have (when taken together with appropriate well-confirmed auxiliary hypotheses) any deductive observational predictions about the results of the experiments in question! In the example we have been considering, the appropriateness of the experimental test in question does not depend on the theory *T*'s having any observational deductive predictions about the results of the experiment. All that is required is that it be *theoretically* plausible that a test of the sensitivity of the relevant biochemical systems to the specific chemical agent

will provide an indication of which of the two accounts of reaction mechanisms (if either) is right. We may test *T* by pitting a hunch about the outcome of experimentation which is theoretically plausible given *T* (and the body of received theories) against an experimental hunch which is theoretically plausible on the assumption of the rival conception of reaction mechanisms. Even though we have assumed that *T* makes a significant number of deductive observational predictions, we need not assume that it makes any *deductive* predictions about the outcome of this crucial experimental test! In sciences which deal with complex systems, instances of theory testing which fit the model just presented are by no means uncommon. Indeed, it may be a good idea to ask whether in describing the instrumental application of theories (rather than their confirmation)—when defining empirical adequacy, for example—the idealization that it is the *deductive* observational consequences of a theory (together with auxiliary hypotheses) rather than its *inductive* consequences that are relevant may not be fundamentally misleading; but that is a topic for another paper.

In any event, what we may learn from these examples is that, in practice, inductive inferences in science extend to inferences with theoretical premises and theoretical conclusions. Just as there are theory-dependent judgments about which possible patterns in observables are projectable, so there are judgments about which patterns in the properties or behavior of "theoretical entities" are projectable. Just as there are theory-dependent judgments of the "degree of confirmation" of instrumental claims by empirical data, so there are theory-dependent judgments of the plausibility of various theoretical claims in the light of other considerations both empirical and theoretical. Indeed, whatever the correct philosophical analysis of this matter, scientific methodology does not dictate any significant distinction between inductive inferences about observables and what certainly look like inductive inferences about unobservables. Finally, and most strikingly, the very methodological principles which govern scientific induction about observables are, in practice, parasitic upon "inductive" inferences about unobservables.

An argument for scientific realism

It will be evident how one may argue for scientific realism on the basis of the theory-dependence of experimental methodology. Consider the question, why are the methodological practices of science instrumentally reliable? Both scientific realists and (almost all) empiricists agree that these practices are instrumentally reliable, but they differ sharply in their capacity to explain this reliability. So theory-dependent are the most basic principles for the assessment of experimental evidence that it must be concluded that these are principles for applying the knowledge which is reflected in currently accepted theories as a guide to the proper methods for the evidential assessment of new theoretical proposals; any other conclusion makes the instrumental success of the scientific method a miracle.

According to the empiricist, the knowledge reflected in the existing body of accepted theories at any time in the history of science is entirely instrumental knowledge: the most we know on the basis of experimental evidence is that the existing body of theories is empirically adequate. Thus, the replacement of existing theories by an empirically equivalent set of theories would leave the knowledge they embody unchanged. Thus, the empiricist can explain the epistemic adequacy of only those theory-dependent features of scientific methodology whose dictates are preserved under the substitution, for the actual body of accepted theories, of any other empirically equivalent one. But, as we have just seen, *none* of the central methodological principles which govern the evaluation of scientific evidence have this property! The consistent empiricist cannot explain the instrumental reliability of the methodology which scientists actually employ.

The scientific realist, on the other hand, has no difficulty in providing the required explanation. According to the realist, existing theories provide approximate knowledge not only of relations between observables, but also of the unobservable structures which underlie observable phenomena. In applying theory-dependent evidential standards, scientists use existing theoretical (and observational) knowledge as a guide to the articulation and experimental assessment of new theories. The judgments of projectability, identification of experimental artifacts, and theoretical criticisms of proposed theories which look ever so much like inductive inferences *are* inductive inferences from acquired theoretical knowledge to new theoretical conclusions. When a theoretical proposal is theoretically plausible in the light of the existing theoretical tradition, what that means is that it is supported by an inductive inference at the theoretical level from previously acquired theoretical knowledge.

Judgments of "projectability" are thus just what they look like "preanalytically": they represent the identification of theoretical proposals for which there are good inductive reasons to believe that they are (approximately) true and thus for which there is good reason to believe that they will eventually be articulated into empirically adequate theories. The role of experimentation is to choose between the various theoretical proposals which pass this preliminary test for probable (approximate) truth.

Similarly, the judgments of theoretical plausibility by which possible experimental artifacts are identified turn out to be inductive inferences from theoretical knowledge which result in reliable assessments of the evidential likelihood that various unobservable factors will influence the outcome of experiments. Finally, the methodological solution to the problem of sampling really does consist in identifying—by reliable inductive inference from theoretical knowledge—the most plausible rivals to a proposed theory and the experimental conditions under which they can be effectively pitted against it. The reliability of scientific methodology in guiding induction about observables turns out to be largely parasitic upon the reliability of the methodology in applying existing theoretical knowledge to guide the establishment of new theoretical knowledge (see Boyd 1973, 1981, 1983). Only this explanation, the realist maintains, can account both for the reliability of the scientific method and for the fact that seemingly inductive reasoning about theoretical matters is so central to it.

Advantages of this defense of scientific realism

The argument for scientific realism just presented has two distinctive advantages when compared with other arguments for realism which emphasize the theory-dependence of scientific practice. In the first place, the standard empiricist response to indications that the methods of science are theory-dependent is to invoke the distinction between the "context of discovery" or "context of invention" on the one hand and the "context of justification" or "context of confirmation" on the other. Theory-dependent features of scientific practice—at least those which depend on the theoretical structure of received theories—are said to be "merely heuristic" features of science which form part of the context of discovery or invention, but which are largely irrelevant to (rationally reconstructed) conformation or justification. The present argument is designed to block such a response by arguing that the relevant theory-dependent features of method are absolutely central to the sort of justification and confirmation which even the empiricist must accept if she holds that scientists can obtain (approximate) instrumental knowledge.

The second advantage to the defense of realism we are discussing is that it provides an epistemologically coherent rebuttal to the empiricist principle that empirically equivalent theories are equally supported or refuted by any body of observations. The evidence for or against a theory is *not* just a matter of the accuracy of its tested empirical predictions; considerations of the theory's own theoretical plausability in the light of received theories, and of the theoretical plausibility of various possible rivals, are essential to its epistemic assessment. The fact (if it is a fact) that a proposed theory (or one of its rivals) is theoretically plausible constitutes inductive warrant for the belief that the theory (or the rival) is (approximately) true. Moreover, the view that considerations of theoretical plausibility are evidential in this way *does not* constitute an abandonment of the doctrine that scientific knowledge is grounded in experiment. The background theories with respect to which theoretical plausibility is assessed are, after all, not *a priori* truths; they have themselves been previously tested by experiment. The theoretical plausibility of a proposed theory thus represents theory-mediated experimental evidence in its favor. Indeed, as we have seen in the discussion of the assessment of experimental evidence, all evidential support which a theory receives from experimental evidence is strongly theory-mediated, even when the evidence involves the confirmation of deductive observational predictions from the theory itself. The empiricist is right that all scientific knowledge is experimental knowledge, but the empiricist conception of experimental evidence fails to include an account of methodologically crucial inductive inferences at the theoretical level; when these are taken into account, the doctrine of the evidential indistinguishability of empirically equivalent theories is evidently false. (For a more elaborate treatment of these issues and of their relation to naturalistic epistemology, see Boyd 1979, 1981, 1983.)

Theoretical inductions and the "unity of science"

Several philosophers have suggested that the so-called "unity of science principle," according to which independently well confirmed theories can be expected to be conjointly empirically adequate, provides a basis for an argument for realism (see, for example, Putnam 1975, chap. 2; Putnam 1978, pt. 3; Boyd 1979, 1981, 1983). The idea here is that, if T and T' are independently tested theories (in the sense that neither has been employed as an "auxiliary hypothesis" in the testing of the other), then the empiricist ought to hold that all that has been tested is the empirical adequacy of each of these theories (together with the auxiliary hypotheses which were used with each in its confirmation). From just the claim that each of the theories is empirically adequate in this sense, it certainly does not follow (even inductively) that the conjunction of the theories will be (even approximately) empirically adequate. Only in the case in which the theoretical vocabularies of the two theories are disjoint would the inference be sound. But the unity of science principle holds (correctly) that the evidence for the two independently tested theories *does* constitute evidence for their conjoint reliability, provided only that no term (theoretical or observational) occurs nonunivocally when the two theories are conjoined.

For the reliability of this principle, the realist offers the explanation that experimental evidence for T and for T' constitutes evidence for the truth of each of the theories and that the judgments of univocality for theoretical terms which are presupposed in the unity of science principle constitute reliable judgments of sameness of reference for the theoretical terms common to T and T'. Thus, under the circumstances envisioned for the application of the unity of science principle, we have evidence that T and T' are true theories with relevantly the same subject matter, and, thus, we have evidence that their conjunction is true. It is this consideration which justifies us in taking the empirical adequacy of their conjunction to be confirmed. (For a more careful statement of this argument, see Boyd 1981,

sec. 2.4.) Here again realism explains the instrumental reliability of a methodological principle whose reliability the empiricist cannot explain.

It is fairly easy to say why the unity of science principle poses a problem for the antirealist. The fact that scientists accept something like the deductive closure of the various individual theories they accept, together with the fact that many of these theories share common theoretical terms, means that in scientific practice (and, in particular, in scientific inductions about observables) theories are integrated in ways which do not appear to make sense unless theories are taken to reflect knowledge of the entities to which their theoretical terms refer.

In the light of the preceding discussion of the role of inductive inferences at the theoretical level in the assessment of experimental evidence, we may offer an even more powerful version of the argument for realism from the unity of science. Not only does sound scientific methodology dictate the *deductive* integration of theories described by the positivists' unity of science principle, it also dictates the *inductive* integration of theories—the use of individually well confirmed theories (sharing common theoretical terms) as premises in inductive as well as deductive inferences. It is just such inferences which are methodologically crucial in the assessment of experimental evidence, and—as we have seen—these inferences make epistemic sense only if the evidence for particular theories is taken to be evidence for their approximate truth (and if our judgments of univocality for theoretical terms are reliable)—that is, only if a realist conception of scientific inquiry is adopted.

It is thus clear that reflection on the crucial epistemic role of inductive inferences at the theoretical level indicates that scientific theories are even more tightly integrated methodologically than the deductive version of the unity of science principle suggests, and the instrumental reliability of the inductive integration of theories provides even greater evidence for a realist conception of theoretical knowledge and a realist conception of the referential semantics for theoretical terms. (For a suitable realist account of reference, see Boyd 1979, 1981.) If the truth be known, the real unity of science is inductive unity, and inductive unity is no respecter of the observation-theory dichotomy.

II. Constructive Empiricism and the Theory-Dependence of Experimental Design

In the present part of this paper, I want to consider the ways in which the empiricist philosopher of science might reply to the arguments in the preceeding sections. I will take van Fraassen 1980 as a paradigm presentation of the empiricist position, but I will try to focus on features of van Fraassen's views which would be broadly acceptable to empiricists, and I will explore some options open to the empiricist which van Fraassen does not explicitly take. A methodological point is in order here: I will not exactly be speculating about the sorts of replies which van Fraassen might offer to the realist arguments just presented. The argument for realism based on the theory-dependence of experimental design is one which van Fraassen discusses in some detail, and so is the argument for realism based on the deductive version of the unity of science principle. It is nevertheless true that I have had the advantage of reformulating the realist arguments in question after having read van Fraassen's thoughtful and stimulating criticisms of them. Since the method of philosophy is dialectical, it can hardly be expected that my account of how the empiricist might reply to these arguments for realism will be definitive. No doubt van Fraassen and other empiricists will have extremely interesting things to say which are more ingenious than those things I say on their behalf. In the meantime, I hope not to be unfair in my treatment of the arguments which van Fraassen has already presented.

I think that we may fruitfully classify van Fraassen's responses to the arguments for realism into five categories: (1) van Fraassen's explicit reply to the argument for realism based on an examination of the methodological principles which govern experimental design; (2) his explicit reply to the argument for realism based on consideration of the unity of science principle in its deductive version; (3) van Fraassen's appeal to a "Darwinist" explanation of the reliability of the experimental method; (4) his discussion of the role in scientific methodology of "pragmatic" virtues of theories like simplicity and explanatory power; and, finally, (5) van Fraassen's discussion of the quite specific philosophical appeal to explanatory power which is involved in the contention that we should believe scientific realism because only scientific realism adequately explains the instrumental reliability of the scientific method.

Van Fraassen on the theory-dependence of experimental design

Van Fraassen discusses at some length the theory-dependence of experimental design and offers an explicit reply to the resulting argument for realism, responding to the version of that argument presented in Boyd 1973. So effective is van Fraassen in his description of the depth to which the scientific standards of experimental design are theory-dependent that I can't resist the temptation to quote him:

> The real importance of theory, to the working scientist, is that it is a factor in experimental design.
>
> This is quite the reverse of the picture drawn by traditional philosophy of science. In that picture, everything is subordinate to the aim of knowing the structure of the world. The central activity is therefore the construction of theories that describe this structure. Experiments are then designed to test these theories, to see if they should be admitted to the office of truth-bearers, contributing to our world-picture.
>
> Whatever the core of truth in that picture (and surely it has some truth to it) it contrasts sharply with the activity Kuhn has termed "normal science", and even with much of what is revolutionary. Scientists aim to discover facts about the world—about the regularities in the observable part of the world. To discover these, one needs experimentation as opposed to reason and reflection. But those regularities are exceedingly subtle and complex, so experimental design is exceedingly difficult. Hence the need for the construction of theories, and for appeal to previously constructed theories to guide the experimental inquiry. . . .
>
> . . . For theory construction, experimentation has a twofold significance: testing for empirical adequacy of the theory as developed so far, *and* filling in the blanks, that is, guiding the continuation of the construction, or the completion, of the theory. Likewise, theory has a twofold role in experimentation: formulation of the questions to be answered in a systematic and compendious fashion, *and* as a guiding factor in the design of the experiments to answer those questions. In all this we can cogently maintain that the aim is to obtain the empirical information conveyed by the assertion that a theory is or is not empirically adequate. (Van Fraassen 1980, 73–74; all subsequent citations are from van Fraassen 1980 unless otherwise indicated.)

When van Fraassen turns his attention to the argument for realism based on the theory-dependence of experimental design, he deals explicitly with a case (from Boyd 1973) of the theory-dependence of what I have here called the solution to the problem of sampling. In that example, a theory L about the mechanisms by which an antibiotic affects a particular bacterial species is proposed and the experimental tests of it which are crucial are identified by considering the sorts of mechanisms which are said by previously accepted theories to operate in, or to interfere with, other antibody-bacteria systems which are (from a theoretical perspective) relevantly similar. The relevant experiments test predictions of L about the effects over time of various doses of the antibiotic on the density of bacterial populations, and the role of theory-dependent considerations is to indicate which of these predictions are most likely to go wrong. The realist argument is that only on a realist understanding both of the theory L and of the knowledge embodied in the relevant background theories can one explain the contribution of these theory-dependent considerations to the reliable testing of L. Here is van Fraassen's rebuttal:

> We must admit that this is one explanation: that the collateral theories are believed to be true. But Boyd needs to establish not only that, as a realist, he can explain what is happening, but also that competing explanations are not feasible.
>
> Let us see then, on Boyd's behalf, how an empiricist can render this methodology intelligible. In the above examples, the collateral theories suggested

> ways in which the function governing population decrease, in terms of drug dosage and time elapsed, might prove to be observably false. Boyd's point is no doubt that the manner in which those theories suggested these consequences, was by suggesting alternative underlying mechanisms which are not directly observable.
>
> I would put this as follows: the models of *L* are quite simple, and reflection on the models of the collateral theories suggests ways in which the models of *L* could be altered in various ways. The empirical adequacy of *L* requires that the phenomena (bacterial population size and its variation) can be fitted into some of its models. Certain phenomena do fit the suggested altered models and not the models of *L* as it stands. Thus a test is devised that will *favour L* (or not favour it) *as against one of those contemplated alternatives*. But it is easy to see that what such a test will do is to speak for (or against) the empirical adequacy of *L* in those respects in which it differs from those alternatives.
>
> The talk of underlying causal mechanisms can be construed therefore as talk about the internal structure of the models. In contrast with the logical, syntactic construal of theories which Boyd used in the discussion of what he called principle 1, we must direct our attention to the family of models of the theory to make sense of the pursuit of empirical adequacy through total immersion (for practical purposes) in the theoretical world-picture. (P. 80)

I take it that van Fraassen is here making a very important point: Once the empiricist abandons the verifiability theory of meaning, it is prefectly possible for her to acknowledge that background theories *suggest* new possibilities to scientists in virtue of the theoretical claims which they make and to describe the reasoning involved as inductive inference from theoretical premises. Acknowledging that scientists make theoretical inductions does not, however, commit the empiricist to holding that the premises of such inductions are possible objects of (noninstrumental) knowledge. ". . . [I]mmersion in the theoretical world-picture does not preclude 'bracketing' its ontological implications" (p. 81). Thus, if the realist were to argue for realism merely on the basis of the fact that scientific practice involves genuine inductive inferences at the theoretical level, she would perhaps refute the verifiability theory of meaning, according to which purely theoretical premises are meaningless, but she would not thereby refute the sophisticated empiricist position advocated by van Fraassen. Quite so. But the problem raised by Boyd 1973 is a different one.

What the realist asks for is an explanation of the contribution of theoretical inductions to the identification of the appropriate experimental tests for proposed theories. The problem is not how background theories can *suggest* to us alternatives to proposed theories or how we can design experiments which test proposed theories against the suggested alternatives. The problem is why the alternatives suggested in this way have a privileged epistemic status, why it is against them and not against other logically possible alternatives that a proposed theory must be tested if it is to receive significant evidential support. Experimental design is "exceedingly difficult" as van Fraassen says, but it would be substantially less difficult if one could legitimately test proposed theories against just any logically possible alternatives. One could then solve the problem of sampling by asking a philosopher to suggest alternatives to the predictions which follow from the theory to be tested (since philosophers are experts on logical possibility). Instead, what good scientific practice requires is that one accept the suggestions which follow by induction from the accepted body of theories. What the empiricist apparently cannot do is to explain why it is *this* solution to the problem of sampling which is instrumentally reliable.

It is unclear whether or not van Fraassen's rebuttal to Boyd 1973 is really intended to address this issue. If so, the best reconstruction of the rebuttal which I can see is this: The relevant background theories are well confirmed, and so there is reason to think that they are empirically adequate. The rivals to *L* which they suggest are, of course, *similar* to these empirically adequate theories, so there are inductive reasons to think that they, too, are empirically adequate. If any one of them is empirically adequate, then *L* is not, so it is inductively reasonable to use such theories as a guide to the circumstances in which *L* is most likely to go wrong (empirically), if it's to go wrong at all. All that's involved is the inductive inference: Such-and-such background theories are empirically adequate; so-and-so alternatives to *L* are similar to these empirically adequate theories; therefore, it is reasonable to believe that one of them is empirically adequate and (therefore) that L makes a false prediction in one or more of the cases in which its predictions differ from those of the so-and-so theories. This is, so far as I can see, exactly right; but, of course, the respects of similarity between the relevant background theories and the suggested alternatives to *L* lie in the theoretically relevant similarities between the accounts they offer of unobservable phenomena (that's how the examples are constructed). The problem for the antirealist is then why these theory-determined respects of similarity are (out of the infinitely many possible respects of similarity) the relevant ones. Nothing van Fraassen says provides an antirealist answer to the basic question of Boyd 1973:

> Suppose you always "guess" where theories are most likely to go wrong experimentally by asking where they are most likely to be false as accounts of causal relations, given the assumption that currently accepted laws represent probable causal knowledge. And suppose your guessing procedure works—that theories really are most likely to go wrong—to yield false experimental predictions—just where a realist would expect them to. And suppose that these guesses are so good that they are central to the success of experimental method. What explanation beside scientific realism is possible? (Boyd 1973, 12)

Van Fraassen on unity of science

Van Fraassen explicitly discusses the argument for scientific realism based on the deductive version of the unity of science principle, which he calls "The Conjunction Objection" (pp. 83–87). It is somewhat difficult to diagnose all of the dimensions of his criticisms of the argument, but two claims stand out. In the first place, van Fraassen seems to claim that *even as an idealization* the deductive version of the unity of science principle does not really describe the epistemic judgments dictated by sound methodology:

> [A]s long as we are scientific in spirit, we cannot become dogmatic about even those theories which we whole-heartedly believe to be true. Hence a scientist must always, even if tacitly, reason *at least* as follows in such a case: if I believe T and T' to be true, then I also believe that (T and T') is true, and hence that it is empirically adequate. But in this new area of application T and T' are genuinely being used in conjunction; therefore, I will have a chance to see whether (T and T') really is empirically adequate, as I believe. *That belief is not supported yet to the extent that my previous evidence supports the claims of empirical adequacy for T and T' separately, even though that support has been as good as I could wish it to be.* [Emphasis mine] Thus my beliefs are about to be put to a more stringent test in this joint application than they have ever been before. (P. 85)

It is true that applications of the unity of science principle do put the conjoined theory to an additional test (as do all previously untried applications of any theory). Therefore, *in a sense*, the as yet untested conjunction of T and T' is not as well supported as the two theories are separately, but it would be a mistake to conclude from this fact that the conjunction in question is evidentially dubious prior to the "more stringent" tests which result from its application. It is well to remember that all cases of theory-determined improvement of measurement and instrumentation in science represent applications of the unity of science principle (Boyd 1981, secs. 2.1, 2.4). Such improvements in instrumentation do not, in the typical case, represent especially speculative or methodologically dubious features of the scientific method; instead, they are central to the improvement of scientific knowledge. The realist has no difficulty in explaining the sense in which successful applications of the conjunction of T and T' represent a more stringent test; what the antirealist apparently cannot explain is why the independent confirmation of T and T' should constitute any significant evidence for the empirical adequacy of their conjunction *at all*! Van Fraassen's first response, although perfectly true, does not address the crucial epistemological challenge raised by the realist.

The other response to the argument for realism from the unity of science which is clear in van Fraassen's discussion involves his insistence that, before scientists accept the conjunction of two well-confirmed theories, they often *correct* them first. As Demopoulos (1982) observes, no realist denies that this happens, and the challenge for the nonrealist to explain the epistemic legitimacy of the resulting integration of independently established theories still remains. I believe, however, that van Fraassen has

nevertheless made a very significant point about the nature of the integration of theories. It may well be that, for many cases involving "small" theories whose adoption does not "make waves," the unity of science principle holds in the strong sense that the conjunction of two well-confirmed theories can be accepted on the basis of the independent evidence supporting each. But, in cases in which the adoption of a particular theory has a serious effect on scientific understanding (and this need by no means entail a "scientific revolution"), some correction of previously accepted theories is often called for, and the necessity of such corrections is often most evident in the case where previously accepted theories are to be applied conjointly with the newly discovered theory. When theories make waves, the ripple effect requires the modification of previously adopted theories.

What is important for the debate between realists and nonrealists is that the modifications in question are themselves theory-determined. From the infinitely many possible modifications of current theories which might be occasioned by the adoption of a new theory (and by whatever new data support it), scientists choose to consider those which are theoretically plausible—which are suggested by inductive inferences at the theoretical (as well as the observational) level. It is the *inductive* integration of theories which is reflected in the sorts of theory modifications which follow a major theoretical innovation. What van Fraassen's insistence on the role of proper correction in the formulation of the unity of science principle correctly indicates is that the unity of science is inductive unity; but, as we have already seen, the epistemic appropriateness of the inductive unity of science is something which only a realist can explain.

Darwinism and the methods of science

Toward the end of chapter 2, after discussing the argument that only realism makes the success of science nonmiraculous, van Fraassen offers a charmingly succinct alternative to the realist's explanation for the success of scientific practice:

> Well, let us accept for now this demand for a scientific explanation of the success of science. Let us also resist construing it as merely a restatement of Smart's "cosmic coincidence" argument, and view it instead as the question why we have successful scientific theories at all. Will this realist explanation with the Scholastic look be a scientifically acceptable answer? I would like to point out that science is a biological phenomenon, an activity by one kind of organism which facilitates its interaction with the environment. And this makes me think that a very different kind of scientific explanation is required.
>
> I can best make the point by contrasting two accounts of the mouse who runs from its enemy, the cat. St. Augustine already remarked on this phenomenon, and provided an intensional explanation: the mouse *perceives* that the cat is its enemy, hence the mouse runs. What is postulated here is the "adequacy" of the mouse's thought to the order of nature: the relation of enmity is correctly

> reflected in his mind. But the Darwinist says: Do not ask why the *mouse* runs from its enemy. Species which did not cope with their natural enemies no longer exist. That is why there are only ones who do.
>
> In just the same way, I claim that the success of current scientific theories is no miracle. It is not even surprising to the scientific (Darwinist) mind. For any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive—the ones which *in fact* latched on to actual regularities in nature. (Pp. 39–40)

In order to assess the plausibility of this alternative explanation, we need to understand more about what biological Darwinism accomplished and to explore the relation between biological Darwinism and van Fraassen's methodological Darwinism. What Darwin and later Darwinists did was to show that the observed adaptation of organisms to their respective environments can be satisfactorily explained without postulating purposive or teleological forces in nature. What van Fraassen's methodological Darwinism aims to show is that the observed adaptation of theories to observable phenomena (i.e., the instrumental reliability of scientific method) can be explained without postulating theoretical knowledge. In the case of biological Darwinism, research attention has, from Darwin 1859 to the present, focused on the *details* of the *mechanisms* of evolution: the mechanisms which produce and constrain heritable variations in phenotypic traits and the various mechanisms—individual selection, sexual selection, kin selection, (perhaps) group selection, genetic "drift", etc.—which produce changes in gene frequency and which are involved in speciation. Thus, to cite two examples among very many, quite special selection mechanisms (kin selection) and quite particular and unusual features of genetic mechanisms (haplodiploidy) are crucial in prominent evolutionary accounts of the evolution of social castes in the Hymenoptera (see Hamilton 1964), and the most significant current debate in the foundations of evolutionary theory is about whether the integration in the 1930s of Mendelian genetics and Darwin's specifically gradualist conception of natural selection (the "modern synthesis") is compatible with fossil data regarding the tempo of evolution and with what is now known about the genetic mechanisms of speciation (see Gould and Eldredge 1977).

What is important for our discussion is that, not only do evolutionary theorists study the particular mechanisms which underlie heritable variation, selection, and speciation, but the success of Darwin's antiteleological project depends upon the results of such studies. As Darwin was acutely aware, there are a number of possible ways in which the mechanisms of heritable variations or of selection might conceivably work such that, if they actually obtained, the evidence would favor a teleological conception of evolution rather than a materialist conception. There are three significant ways in which facts about genetic variation and evolutionary selection

mechanisms *might* have undermined Darwin's antiteleological position. For each of these there is an analogous way in which facts about variation and selection in the evolution of scientific theories might tell against the antirealist. If we follow the example of evolutionary biologists and examine the details of the mechanisms of the evolution of theories, then, I believe, we will find that van Fraassen's antirealist conception of the evolution of theories is refuted at just the points suggested by analogy with the ways in which Darwin's antiteleological position *might* have been refuted. I turn now to a consideration of those points:

1. *Directed variation.* The heritable variations upon which the mechanisms of natural selection operate are limited: not all conceivable phenotypic variations occur, and not all of those which occur are heritable. In general, limitation on heritable variation reduces the effectiveness of natural selection (or kin selection, etc.) in producing adaptations of organisms to their environments. Darwin emphasizes that evolution is opportunistic, "seizing" the opportunities provided by actually occurring variations rather than achieving the perfection in design which might result if heritable variations in the relevant traits were less limited. It *could be*, however, that the natural constraints on heritable variation contribute to, rather than diminish, the effectiveness of evolution in producing adaptations. Heritable variations might be "directed" in the sense that variations are more likely to occur when they would be useful to the relevant organisms. Were this the case, there would be *prima facie* evidence for the operation of teleological mechanisms in evolution, and Darwin is at pains to deny that heritable variation is typically directed in this sense. He proposes that all actual cases of directed variation can be explained by the inheritance of acquired characteristics and other similar mechanisms of a nonteleological sort. If facts had been other than they are—if biological limitations on heritable variation had turned out to contribute to the effectiveness of selection, and if no mechanism like inheritance of acquired characteristics had been sufficient to explain this phenomenon—then there would have been good evidence for the existence of teleological factors in nature, and Darwin's materialist position on this issue would have been *prima facie* refuted.

What is the analogous issue for the antirealist conception of the evoluion of theories? The analogue in the case of the evolution of theories to the biological limitations on heritable variation are whatever methodological limitations act to constrain the range of possible theories which are subjected to "selection," that is, to experimental testing and to other sorts of methodologically appropriate evaluation. The analogue of the issue of directed variation is the issue of whether these methodological limitations act to increase the effectiveness of selection—that is, to increase the instrumental reliability of scientific methodology—in ways which cannot be explained without postulating theoretical knowledge. As we have seen in section 1, the methodological limitations in question are just those which represent the methodological solution to the generalized problem of "projectability" for patterns in observable data, and indeed these limitations *do* contribute to the instrumental reliability of scientific methods in a way which cannot be explained in a nonrealist fashion.

2. *The efficiency of selection.* Darwin 1859 introduces the idea of natural selection by considering variation and selection of organisms under domestication. The purposive selection involved in careful breeding of plants and animals, Darwin argues, can lead to the establishment of strains which are so phenotypically different that they appear to represent different species or even different genera and which are "adapted" to the needs and interests of breeders. Given the evidence for the great antiquity of fossil organisms and of the earth, Darwin suggests, it is reasonable to hold that the nonpurposive and much less efficient mechanisms of natural selection could have produced even more substantial changes and even more various and refined adaptations (in this case, to the demands of the environment). Even if we assume that Darwin was right that species originated by selection operating on heritable variations, and even if we assume that he was right that heritable variations are not teleologically determined, it remains true that Darwin's antiteleological position would have been refuted if the data had shown that selection had been *too* efficient to be explained by any plausible materialist theory. Suppose, for example, that new data about the age of the earth or about the actual rate of speciation, together with findings about genetic mechanisms and about the effects of various natural mechanisms of selection, had made it impossible to explain the extent of adaptation or of speciation merely in terms of the interaction of the relevant genetic mechanisms with naturally occurring mechanisms of selection. In that case, the evidence would have favored a teleological conception according to which something rather like artificial selection guided biological evolution. Darwin's position would have been (*prima facie*) refuted by an observed efficiency of evolutionary selection mechanisms too great to be explained in materialist terms.

What is the analogous issue for the antirealist Darwinian conception of the evolution of theories? The analogues to biological selection mechanisms are the various mechanisms for theory evaluation which are dictated by scientific method. The analogue to the issue of efficiency of the mechanisms of biological selection is the issue of whether or not the important contribution of the methodological principles which govern theory testing to the instrumental reliability (efficiency) of scientific practice can be explained without postulating theoretical knowledge. But these are just the theory-dependent mechanisms by which the problems of experimental artifacts and of "sampling" are solved, and we have already seen that only a realist can explain their contribution to the instrumental reliability of scientific methods. Once again, if we follow the example of biologists and examine closely the mechanisms of theory selection, we see that the antirealist view of the evolution of theories is not sustained by the data.

3. *Future-directed selection.* It is a consequence of any Darwinian conception of evolution that, insofar as natural selection tends to establish adaptive traits, the relevant measure of adaptiveness is the reproductive fitness of the particular organisms which have the traits in question *rather than* the long-term survivability of the species. It is commonplace, for example, to cite cases of selection for fecundity which eventually helps to create conditions of overpopulation leading to extinction. The conception that selection cannot be directed toward the long-term needs of species except insofar as those needs are captured by present reproductive needs is, of course, *prima facie* dictated by a materialist and antiteleological conception of evolution. If it had turned out that evolutionary mechanisms tended to establish traits which reduced immediate reproductive fitness but which served instead the long-run interest of species survival, then, barring some materialist account of this remarkable phenomenon, there would have been strong *prima facie* evidence for a teleological conception of evolution and against Darwin's conception.

Is there an analogous issue for the theory of the evolution of scientific theories? I think the issue is this: Do the actual mechanisms of theory evaluation permit us to assess the *future* empirical reliability of the theories we test (*future* in the sense of going beyond the reliability actually demonstrated in the outcomes of the relevant experiments) in a way which cannot be explained without postulating theoretical knowledge? Here again, the answer discovered by a careful examination of the methods of science refutes the nonrealist interpretation of the evolution of scientific theories. Scientific methodology does indeed allow us to assess the future empirical reliability of theories by solving the problems of projectability, experimental artifacts, and sampling, but the ways in which it does so are inexplicable except from a realist perspective. The unity of science principle illustrates what is essentially the same phenomenon. The independent testing of particular scientific theories serves to establish their approximate instrumental reliability with respect to both deductive and inductive applications in future (conjoint) contexts quite unlike those in which their reliability has been "directly" tested. Here, too, the only explanation for this future-regarding feature of the methodological selection of theories is one which postulates theoretical knowledge.

I conclude that van Fraassen's appeal to the analogy between theory testing and natural selection provides a valuable perspective from which to study the epistemology of science but that—when we carry out the scientific investigations suggested by the analogy—it turns out that a realist rather than an antirealist conclusion is dictated by the available evidence.

There is one more lesson to be learned from the analogy with evolutionary theory. Van Fraassen's antirealist appeal to a Darwinian conception of theory evolution is strikingly reminiscent of an approach to biological evolution called (by its followers) "optimization theory" or (by its critics) "adaptationism." Roughly, the idea of this approach is that, relatively independently of considerations of the details of evolutionary and genetic mechanisms, one may generally expect natural selection to produce near optimal "solutions" to the "problems" set for organisms by their environments. If this is so, then within certain limits one can take the problem-solving capacity of natural selection for granted without examining the details of the underlying mechanisms. By analogy, one might think it appropriate in epistemological contexts to take the instrumental reliability of scientific methodology for granted, subsuming it under the broad category "induction" and taking the notion of induction pretty much for granted. Gould and Lewontin (1979) have offered serious criticisms of adaptationism in evolutionary theory. Whatever may be the merits of their criticisms, there seems little doubt that analogous arguments of Goodman (1973) show that one cannot take the foundations of induction for granted in epistemology.

> Consider the question: How is reasonable expectation about future events possible? ("Future" may be replaced by "unobserved" for generality.) The recurrent idea that there is some rational form of simple extrapolation from the past, something like rules of induction, may be especially appealing to empiricists because it holds out hope for a presuppositionless, non-metaphysical answer. But it is an idea that goes into bankruptcy with every new philosophical generation. (Van Fraassen 1982, 26)

Van Fraassen on pragmatic virtues

According to the realist conception of scientific knowledge defended in section 1, the fact that a proposed theory is inductively supported at the theoretical level on the basis of already confirmed theories constitutes (some) evidence in favor of its approximate truth. The sort of inductive support at issue is often described in the literature in terms of explanatory power. Typically, when there are good theoretical reasons to believe a theory, then there are good reasons to believe that it provides the right explanation for whatever phenomena it describes; typically, when a theory is said to provide a good (or the best) explanation of some observable phenomenon, what is being reported is that there are *theoretical* reasons (and often experimental reasons as well—but these are theory-dependent, as we have seen) to believe that it provides an approximately true account of how those phenomena come about. I have here avoided reference to the principle of "inductive inference to the best explanation" (Harman 1965) largely because I think that this formulation carries the implicature that there is some conception of explanation, like the "common-cause principle" (pp. 25–31; 118–33), which can be established prior to theoretical understanding and which can serve to justify inferences from observations to theoretical conclusions. In the view expressed here, what are important are inductive inferences from (partly or wholly) theoretical premises to (partly or wholly) theoretical conclusions. The "rules" governing such inductive inferences (judgments of projectability for properties of various sorts of "theoretical entities," for example) are themselves theory-determined. There are no significant pretheoretical rules of inductive inference at either the theoretical or the observational level in science. Thus, insofar as inductive theoretical reasoning can often be described in terms of reasoning about explanation, the relevant notions of explanation should themselves be theory-dependent; our theoretical knowledge tells us what sorts of explanations are possible and what standards are to be used to judge them.

Despite the divergence between the conception of theoretical reasons defended here and the conception of inductive inference to the best explanation considered by van Fraassen, it seems reasonable to take what van Fraassen says about explanatory power, simplicity, elegance, capacity for unification of disparate phenomena, and other such virtues of theories as the basis for a possible empiricist reply to the arguments for realism rehearsed in section 1. Van Fraassen calls these virtues the "pragmatic virtues," and, when he characterizes the dimension which they add to theory acceptance, he does indeed seem to have in mind the sort of theoretical-tradition-dependent evaluative considerations on which the argument for realism depends:

> *Theory acceptance has a pragmatic dimension.* While the only belief involved in acceptance, as I see it, is the belief that the theory is empirically adequate, *more than belief is involved.* To accept a theory is to make a commitment, a commitment to the further confrontation of new phenomena within the framework of that theory, a commitment to a research programme, and a wager that all relevant phenomena can be accounted for without giving up that theory. That is why someone who has accepted a certain theory, will henceforth answer questions *ex cathedra*, or at least feel called upon to do so. (P. 88)

Van Fraassen's treatment of the pragmatic virtues is just what the term *pragmatic* suggests:

> In so far as they go beyond consistency, empirical adequacy, and empirical strength, they do not concern the relation between the theory and the world, but rather the use and usefulness of the theory; they provide reasons to prefer the theory independently of questions of truth." (P. 88)

Here, I believe, van Fraassen makes an extremely important point. If we think of scientists as somehow already knowing about certain scientific theories that they are empirically adequate, then it is an uphill battle for the realist to have to argue that the additional fact that these theories are also explanatory, elegant, simple, harmonious with previous theories, etc., makes it likely that they are (even approximately) true. As I have argued elsewhere (Boyd 1981, 1983), the realist who takes it for granted that (s)he and the empiricist are each already able to assess the experimental evidence for and against the empirical adequacy of scientific theories will find it impossible to use the standard arguments for realism against the powerful epistemological arguments which appear to support empiricist anti-realism. Van Fraassen seems well aware of this difficulty for realists when he characterizes as "rock-bottom criteria of minimal acceptability" "consistency, internally and with the facts," while denying that explanatory power is a rock-bottom virtue in this sense (p. 94).

The defense of realism presented in section 1 does not, of course, face these difficulties. The proposal to understand the theory-dependent virtues as merely pragmatic is subject to the same objections raised in section 1 to the treatment of those virtues as "merely heuristic." Either treatment requires that assessments of the "rock-bottom" virtue of empirical adequacy be independent of theoretical knowledge. But they are not. It is only because our *theoretical* commitments reflect approximate knowledge of unobservables that we are able to assess the empirical adequacy of scientific theories. It *may* be true that our knowledge of the observable results of particular experiments is "rock-bottom" in a way that theoretical knowledge is not, but our knowledge about scientific theories that they are empirically adequate is typically parasitic on our knowledge of "theoretical entities."

Kuhn and others have shown that the justifications scientists actually give for the inductive generalizations they make about observables (their justifications for particular inductive strategies, experimental designs, patterns of inference, etc.) are profoundly theory-dependent. What the realist investigation of the problems of experimental design shows is that this situation cannot be remedied by "rational reconstruction": the only good justifications there are for the inductive practices of scientists are theoretical justifications of the sort only a realist can accept. In the light of these findings, the view that theoretical considerations are merely pragmatic or merely heuristic dictates the absurd conclusion that the inductive inferences about observables which scientists make are without justification (see Boyd 1983, especially part 8).

Inductive inference to philosophical explanations

I have just argued that various empiricist rebuttals to the argument for realism indicated in section 1 are inadequate; the empiricist cannot escape

the conclusion that only from a realist perspective is it possible to explain the instrumental reliability of the actual methods of science or the legitimate role of theoretical considerations in those methods. Suppose (as is unlikely) that the empiricist grants this conclusion. What argumentative resources does the empiricist still have at her disposal? One response is clearly suggested by van Fraassen's treatment both of the arguments for realism in particular and of the issue of inductive inference to the best explanation in general: The empiricist must, anyway, reject the epistemic legitimacy of inferences to the best explanation when the conclusions of those inferences are about unobservables; the empiricist may then in particular reject the realist conclusion, even granting that realism provides the best explanation for the instrumental reliability of the methods of science (see also Fine 1984). If the empiricist adopts this response, she will of course still need to retain the doctrine that the methods of science are instrumentally reliable. This is so for two reasons.

First, the doctrine that the methods of science are instrumentally reliable is a generalization *about observables* which is quite evidently true; any philosophical position which abandoned it would be *prima facie* refuted. Secondly, it is the business of philosophers of science, empiricist or otherwise, to identify and discuss the principles which constitute good scientific methodology. The identification of such principles depends on ascertaining which features of scientific practice contribute to the epistemic reliability of science. For the empiricist, the identification of such features will depend precisely on generalizations of the sort in question regarding the contribution of such features to the instrumental reliability of scientific practice. It is also true that, if the empiricist can inductively establish the generalization that the methods of science are instrumentally reliable, then she will have the resources to reply to the claim of the preceding section that the empiricist cannot explain why scientists are justified in making inductive generalizations about observables. The empiricist can offer, on behalf of the scientist, an inductive justification of the inductive practices of science: It is inductively confirmed that the theory-dependent methods of science are instrumentally reliable. Therefore, in any particular case, scientists are justified in accepting as empirically adequate a theory established by those methods, even though they are not justified in adopting any beliefs about unobservables.

I have elsewhere discussed this empiricist strategy in detail (Boyd 1983, pt. 8); I conclude there that the consistent empiricist cannot justify the inference from the observed history of scientific practice to the generalization that the methods of science are instrumentally reliable. I will summarize here one of the arguments which supports this conclusion. In order to establish inductively the instrumental reliability of the scientific method, one must know that it has been appropriately reliable in the past. In order to know this, one must know about various theories which have been accepted by scientists in the past that *they* are approximately empirically adequate. As Kuhn has shown, scientists cannot offer justifications for such conclusions which do not depend upon theoretical beliefs which the empiricist cannot accept. As the argument for realism discussed in section 1 shows, this problem for the empiricist cannot be solved by "rational reconstruction." The inductive justification for theory-dependent inductions about observables cannot be invoked by the empiricist, because the generalization whose justifiability we are discussing is a *premise* for that inductive justification. Therefore, the *consistent* empiricist cannot even justifiably conclude that the methods of science have been instrumentally reliable in the past, much less that they will be reliable in the future.

I conclude that the consistent empiricist can justify neither the methods and empirical findings of science nor the methods and findings of empiricist philosophy of science.

III. Credo: *Lex Orandi est Lex Credendi* (*Analogice Acceptum*)

In chapter 7, "Gentle Polemics," van Fraassen treats the reader to a clever and engaging comparison of various arguments for scientific realism with Aquinas's arguments for the existence of God. The reader, unconvinced by the Thomistic arguments, is supposed to reject the realist's arguments as well. If, as I argued in the previous section, the consistent empiricist cannot justify either the sorts of empirical generalization represented by scientific theories or the basic methods of the philosophy of science, then it is perhaps fair to suggest that the empiricist's position has come to resemble that attributed to Tertullian (160–220), "*Credo quia absurdum.*" In any event, the initial comparison of the realist's arguments with those of Aquinas cannot be sustained in the case of the arguments for realism presented here.

If one looks for precedents for those arguments, they are best found in the persistent "flirtation" with atomism or metaphysical materialism which has characterized the empiricist tradition from Locke to Carnap or in the dialectical materialism of the Marxist tradition. If a theological precedent must be found, the obvious choice is the dictum attributed to Pope St. Celestine I (422–32; the attribution may be faulty—see Buchberger 1954, vol. 6, col. 1001): "*Lex orandi est lex credenti*," "the rule for praying is the rule for believing," or (in a freer translation) "believe what is necessary to 'rationally reconstruct' liturgical practice." For "liturgical practice" put "scientific practice" and you get the strategy for the defense of realism employed here.

Drawing the analogy between this strategy and the principle of theological methodology attributed to Celestine is helpful in understanding the

former. The Celestine methodological principle differs sharply from the inferential principles involved in Aquinas's "proofs" of the existence of God. Aquinas appeals to extremely general principles about explanation and causation which are apparently supposed to be basically *a priori* and to be universally applicable. The Celestine principle, by contrast, is not *a priori* and is (even in the view of its defenders) not universally applicable. It is supposed to apply only with respect to the right sort of liturgy, so that a theist considering employing the principle would not know *a priori* whether or not her application of the principle would be epistemically reliable. The Celestine principle is a methodological principal *within* theology, not prior to it, and it is understood to be reliable only within a particular theological tradition.

The analogous claims are true of the methodological principles underlying the defense of realism discussed here. What the realist proposes is to use the ordinary methods of science to investigate the question of why the methods of science are instrumentally reliable. The philosophical methods here are not conceived of as *prior to* scientific methods in any sense. Moreover, according to the realist's own account, the reliability of the scientific methods in question depends on the approximate truth of the background theories in the theoretical tradition; thus, the reliability of the realist's *philosophical* methods depends on logically and epistemically contingent facts about the actual scientific tradition (Boyd 1981, 1983). The realist's philosophical methods are *in that sense* not universally applicable. The principles of inference by which the realist defends realism will be no more stringent than the principles of inference whose reliability the realist is trying to explain. Indeed, if the realist is right, then the principles of scientific methodology to which the realist appeals are tacitly but ineliminably realist themselves. The realist—in insisting that the methods of the philosophy of science should be the methods of science—cannot not offer to defend realism by appeal to inferential principles which are, in the final analysis, themselves neutral with respect to the realism-empiricism controversy. This may seem a grave defect in the realist's program (see, for example, Fine 1984, fn. 7). The realist is content to reply that it is also impossible to defend the inductive inferences which scientists make about observables if one insists on inferential principles neutral with respect to the realism-empiricism controversy; it is likewise impossible thus to defend the philosophical methods of even empiricist philosophers of science (Boyd 1983). Apparently both in science and in the philosophy of science we must make do with Celestine rather than Thomistic principles. For the philosopher who—in the name of science—objects to the *a priorism* of Thomistic reasoning, this may not be an entirely unwelcome conclusion.

References

Boyd, R. 1973. "Realism, Underdetermination and a Causal Theory of Evidence." *Nous* 7:1–12.

———. 1979. "Metaphor and Theory Change." In *Metaphor and Thought*, edited by A. Ortony, 356–408. Cambridge: Cambridge University Press.

———. 1981. "Scientific Realism and Naturalistic Epistemology." In *PSA 1980*, vol. 2, edited by P. D. Asquith and R. N. Giere, 613–62. East Lansing, Mich.: Philosophy of Science Association.

———. 1983. "On the Current Status of the Issue of Scientific Realism." *Erkenntnis* 17:135–69.

Buchberger, M., ed. 1954. *Lexikon für Theologis und Kirche*. Freiburg.

Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection*. London: John Murray.

Demopoulos, W. 1982. Review of *The Scientific Image*, by Bas C. van Fraassen. *Philosophical Review* 91:603–7.

Fine, A. 1984. "The Natural Ontological Attitude." In *Essays on Scientific Realism*, edited by J. Leplin.

Goodman, N. 1973. *Fact, Fiction and Forecast*. 3d ed. Indianapolis and New York: Bobbs-Merrill.

Gould, S. J., and N. Eldredge. 1977. "Punctuated Equilibria: The Tempo and Mode of Evolution Reconsidered." *Paleobiology*, 1977, 115–51.

Gould, S. J., and R. Lewontin. 1979. "The Spondrels on San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." *Proc. R. Soc. Lond.* 3.205, 581–98.

Hamilton, W. D. 1964. "The Genetic Theory of Social Behavior," I, II. *Journal of Theoretical Biology* 7 (1): 1–52.

Harman, G. 1965. "The Inference to the Best Explanation." *Philosophical Review* 74:88–95.

Kuhn, T. 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Putnam, H. 1975. *Mind, Language and Reality: Philosophical Papers*, vol. 2. Cambridge: Cambridge University Press.

———. 1978. *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.

Van Fraassen, B. 1980. *The Scientific Image*. Oxford: Clarendon Press.

———. 1982. "The Charybdis of Realism: Epistemological Implications of Bell's Inequality." *Synthese* 52:25–38.

2 The Ontological Status of Observables: In Praise of the Superempirical Virtues

Paul M. Churchland

At several points in the reading of van Fraassen's book, I feared I would no longer be a realist by the time I completed it. Fortunately, sheer doxastic inertia has allowed my convictions to survive its searching critique, at least temporarily, and, as we address you today, van Fraassen and I still hold different views. I am a scientific realist, of unorthodox persuasion, and van Fraassen is a constructive empiricist, whose persuasions currently define the doctrine. I assert that global excellence of theory is the ultimate measure of truth and ontology at all levels of cognition, even at the observational level. Van Fraassen asserts that descriptive excellence at the observational level is the only genuine measure of any theory's truth and that one's acceptance of a theory should create no ontological commitments whatever beyond the observational level.

Against his first claim I will maintain that observational excellence or 'empirical adequacy' is only one epistemic virtue among others of equal or comparable importance. And against his second claim I will maintain that the ontological commitments of any theory are wholly blind to the idiosyncratic distinction between what is and what is not humanly observable, and so should be our own ontological commitments. Criticism will be directed primarily at van Fraassen's *selective* skepticism in favor of observable ontologies over unobservable ontologies and against his view that the 'superempirical' theoretical virtues (simplicity, coherence, explanatory power) are merely pragmatic virtues, irrelevant to the estimate of a theory's truth. My aims are not merely critical, however. Scientific realism does need reworking, and there are good reasons for moving it in the direction of van Fraassen's constructive empiricism, as will be discussed in the closing section of this paper. But those reasons do not support the skeptical theses at issue.

Previously published in a shorter form as "The Anti-Realist Epistemology of van Fraassen's *The Scientific Image*," *Pacific Philosophical Quarterly* 63 (July 1982): 226–36. Reproduced by permission. I thank Hartry Field, Michael Stack, Bas van Fraassen, Clark Glymour, Barney Keaney, Stephen Stich, and Patricia Churchland for helpful discussion of the issues here addressed.