THE UNIVERSITY OF CALGARY

Explanatory Inference and Entity Realism

by

Geoffrey Gorham

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF
MASTER OF ARTS

Department of Philosophy

Calgary, Alberta
September, 1988

ISBN 0-315-46578-6

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, "Explanatory Inference and Entity Realism" submitted by Geoffrey Gorham in partial fulfillment of the requirements for the degree Master of Arts.

John W. Heintz

Supervisor, John Heintz
Department of Philosophy

John Collier

John Collier
Department of Philosophy

Margaret J. Osler

Margaret J. Osler
Department of History

September 12, 1988

ABSTRACT

In contemporary philosophy of science debates concerning scientific realism focus primarily on epistemological, rather than metaphysical or semantic, issues. In particular, the most fundamental point of contention between realists and anti-realists is the legitimacy of the principle of inference to the best explanation.

Two anti-realist arguments--the argument from empirical underdetermination and the historical gambit--are fatal to the realist's employment of the principle of inference to the best explanation as a means of justifying belief in the truth of any particular scientific theory currently held. Nevertheless, a second-order version of the principle, aimed at establishing a realist thesis about the general aim of science, does not fall to these objections.

However, this meta-version of the principle of inference to the best explanation can succeed only by assuming a realist theory of explanation, and hence cannot be decisive against anti-realism. Anti-realists like van Fraassen understand explanation as an irreducibly pragmatic feature of scientific theories, distinct from question about their truth or falsity. Hence, for the anti-realist explanatory power furnishes grounds for acceptance, but never for belief.

But this pragmatic view of explanation cannot apply to singular causal explanations since such explanations have an inherent existential component. This leads to the position of theoretical entity realism via inference to the best causal explanation, a narrower version of the standard realist inferential principle. Besides inference

to the best causal explanation, arguments based on experimental intervention are also offered in defense of entity realism. Interventionist arguments ultimately depend for their validity on inference to the best causal explanation. Entity realism suggests a conception of scientific progress that is counter to the unity of science principle.

ACKNOWLEDGEMENTS

I would like thank my supervisor Dr. John Heintz for his philosophical guidance and encouragement, as well as Dr. Margaret J. Osler and Dr. John Collier for their helpful advice and comments towards the completion of this project.

I would also like to thank the members of the philosophy department, particularly my fellow graduate students and friends John Pugsley, Hugh Leighton, Mark Gardiner and Lance Cooke, who made my stay in Calgary a time I will long remember. Special thanks to Damon Gitelman whose careful proof-reading and philosophical insight probably saved this thesis from the depths of obscurity.

Finally I would like to acknowledge in a rather different way the Alberta Minister of Advanced Education, Hon. Dave Russell. His government's draconian educational funding policies during my term of residence at the University of Calgary prompted a year-long involvment in student politics, leaving me less time to research and write this thesis than I might have had. I hope it was worth it.

For my parents, Carl and Donna Gorham, with love and thanks for all their help and guidance.

TABLE OF CONTENTS

INTRODUCTION

Since the demise of logical positivism, scientific realism has re-emerged in a variety of interesting forms. In fact, there have been almost as many versions of scientific realism as there have been scientific realists. But within this multiplicity of realist theses I think it is possible to identify three basic forms of scientific realism. The earliest post-positivist version might be called 'naive scientific realism'. This position, which was defended by Wilfred Sellars and the early Putnam, held simply that the terms of mature scientific theories typically genuinely refer, and that the theories accepted in a mature science are typically approximately true. The second version to become popular was more moderate. This view, which I shall call 'progressive realism', operates at the meta-level of scientific inquiry, focusing on the overall success of scientific methodology, rather than the statements of any particular scientific theory. The 'progressive realist' defends a realist thesis about the aim of science--science is providing an increasingly accurate description of reality at both the observable and unobservable level--on the grounds that only this view will allow for a satisfactory explanation of the increasing empirical success of science through history. Contemporary proponents of 'progressive realism' include Richard Boyd and Ernan McMullin. The third and most recent form of scientific realism to appear on the philosophical scene is 'entity realism'. Entity realists, such as Nancy Cartwright and Ian Hacking, recommend belief in the existence of the concrete entities and physical structures postulated by contemporary scientific theory, but not in the truth of the basic principles and laws of those theories.

This evolution in the character of scientific realism over the last thirty or so years can be explained by the corresponding emergence of several sophisticated critiques of scientific realism from philosophers such as Bas van Fraassen and Larry Laudan. The arguments of these scientific anti-realists call into question the central inferential strategy of both the 'theory-specific' approach of 'naive scientific realism', and the meta-level reasoning of the 'progressive realists'.

When one surveys the arguments of the realists and anti-realists over this period, two basic underlying features of the controversy become apparent. First, the debate has, and continues, to center primarily on epistemological questions. That is, the fundamental issues concern the justification of belief in scientific theories, as opposed to questions about the meaning of theoretical language or the metaphysical status of theoretical postulates. Second, within epistemology, arguments from both sides have concentrated in one way or another on the legitimacy of abductive inferences, and particularly inference to the best explanation. In other words, the most powerful arguments for scientific realism have been explanationist in form. My main thesis in what follows is that the best way to understand the realist - anti-realist controversy in recent philosophy of science is to recognize the centrality of these two points. I think that from this perspective, one can comprehend with the most precision and ease both the failings of earlier forms of scientific realism, and the successes of the reigning king--entity realism.

For these reasons I think this thesis should be understood as having two basic philosophical components. The first component is the attempt--to use Lakatos' famous phrase--at a 'rational reconstruction' of the history of the realist - anti-realist debate in recent years. The second component, which is intimately related to the first, is a defense of entity-realism as the only form of scientific realism which will surive the most persuasive attacks of the anti-realists.

In the first chapter I show in some detail how epistemology has come to occupy a place of central importance in the contemporary debate. I argue that neither metaphysics nor semantics are especially relevant to the main arguments of either realists or anti-realists--in these areas they generally agree. Instead, the opposition lies in opposing attitudes concerning the epistemological significance of abduction and inference to the best explanation.

In the second chapter I present and analyse the two most important arguments that have been offered from the anti-realist camp against any simple theory-specific employment of the principle of inference to the best explanation towards realist conclusions. I argue that, together, the argument from the empirical underdetermination of theory by data and the pessimistic induction from the history of science, undermine any direct inference from the emprical success of scientific theories to their truth. But I argue further that the principle of inference to the best explanation can still be invoked to justify a realist thesis about scientifc progress, without succumbing to these two basic anti-realist objections.

In the third chapter, I introduce a deeper objection to the realist's adherence to inference to the best explanation. I contend that explanatory inference can only support uniquely realist conclusions by assuming a uniquely realist view of explanation. Anti-realists such as van Fraassen see explanation as an essentially pragmatic virtue of scientific theories, which warrants acceptance of theories, but does not provide independent grounds for belief. Hence, the fact that 'progressive realism' is the best explanation for the success of science cannot settle the fundamental debate about the relation between truth and explanation without begging the question against the anti-realist.

In the fourth and final chapter I argue that van Fraassen's pragmatic theory of explanation does not support agnosticism where the scientific explanations in questions are causal. Following Nancy Cartwright, I contend that the acceptance of causal explanations carries ontololgical commitment to the entities appealed to as explanans. This suggests the adoption of the principle of inference to the best *causal* explanation, which constitutes the philosphical basis for entity realism. A different sort of argument for entity realism, which focuses on our ability to manipulate unobservable entities in experimental contexts, has force against anti-realism only if inference to the best causal explanation is legitimate to begin with. I conclude the final chapter with some some reflections on the implications of entity realism for a philosophical account of scientific progress.

It is commonly charged against philosophers of science that they write about their subject matter purely in the abstract, rarely deigning to descend to the concrete level of historical case study and actual scientific activity. This tendency is especially apparent in the work of the logical postivists, though, even with an increasing attention in recent years on the part of philosophers to the history of science, there is still no shortage of philosophical tracts about science which are only remotely about actual scientific practice.

Ideally, philsosophy of science should aim to walk a fine line between abstract analysis on the one hand, and the simple description of scientific activity on the other. It is necessary to maintain a certain degree of intellectual detachment when thinking philosophically about science in order to enable general conclusions about its aim and function. But with too much detachment, the philosopher often will end up merely doing logic, epistemology or metaphysics, in the guise of philosophy of science.

I suspect that the present work unfortunately falls rather heavily on the abstract side of this precarious balance. I offer in my own defence only the following reminder of what was said above. This thesis is primarily about the philsophy of science rather than science itself. In particular it is about the *epistemology* of science, and the degree to which explanatory inference can ground a realist view of scientific inquiry. I think the real danger of philosophical abstraction in writing about science is for those philosophers (such as the positivists) who attempt to identify *the correct* scientific method. Such an enterpise is doomed to irrelevance when undertaken in ignorance of what scientists actually do. It remains though that any work in the philsophy of science will gain by a clear understanding and appreciation of the complexities of the actual activity of science, now and over the course of history. The fact that my thesis does not profit from laboratory experience and a close study of primary historical tracts I attribute to a lack of time.

CHAPTER ONE:
THE EPISTEMOLOGY OF SCIENTIFIC REALISM

Introduction

Contemporary debates concerning scientific realism focus primarily on epistemological issues. They center on questions regarding the significance and legitimacy of certain kinds of inferential strategies employed in scientific contexts. I would hold, in particular, that a good deal of the disagreement between realists and anti-realists has its source in conflicting attitudes towards abductive inference, and especially inference to the best explanation. Abduction is a term originally used by C. S. Peirce to identify that process by which scientists formulate hypotheses to account for phenomena. Formally, abduction begins with a premise describing a given phenomenon. A second premise states how some other state of affairs might be expected to give rise to that phenomenon. The conclusion is that the proposed state of affairs holds. In this way, for example, the existence of Pluto was originally determined abductively to account for the apparent irregularities in planetary motion. While abduction need not lead to the postulation of *unobservable* entities or processes--consider the usual method used by T.V. detectives to discover the identity of the murderer--it is this sort of inference that is relevant to the realist - anti-realist debate.

On the face of it abduction will lead to the assertion of a multitude of incompatible states of affairs, since any number of distinct sorts of entities and processes may be expected to produce a certain kind of phenomenon. Peirce himself realized this, and in fact intended abduction as only the first step towards the adoption of a scientific theory. Abduction serves only to delimit the possible explanatory hypotheses, which are subsequently subjected to the usual experimental techniques.

On the other hand, a special instance of the abductive style, inference to the best explanation, has as its main premise the definite description "__ is the single best explanation for the given phenomenon". Thus, inference to the best explanation combines the logical form of abductive inference with knowledge of the results of experimentation, so to arrive at a unique conclusion. As Gilbert Harman puts it in his original discussion of inference to the best explanation, "One infers from the fact that a given hypothesis would provide a 'better' explanation for the evidence than would any other hypothesis to the conclusion that the given hypothesis is true."[1] While we will see that there are other problems with the principle of inference to the best explanation, it overcomes this initial logical difficulty of deriving incompatible conclusions from the same set of premises.

Anti-realists generally adhere to a principle of minimizing epistemic risk, and hence resist judgments that extend beyond the "safe" realm of the observable world. They therefore reject any purported ontological implications of abductive-style inferences, preferring instead to adopt an agnostic stance towards the microstructural aspects of scientific theories. The general strategy, as we will see later, is to sever any supposed connection between

---

1. Harman (1965), p. 89

explanatory power and truth, thereby rendering inference to the best explanation invalid, and robbing abduction of any realist implications. The classic statement of this form of scientific anti-realism is to be found in Pierre Duhem's The Aim and Structure of Physical Theory, and has its best recent exposition in Bas van Fraassen's The Scientific Image. These authors maintain that the aim of science should be understood as that of providing a true account of the observable--to save the phenomena; questions regarding the nature of the unobservable are undecidable. Inasmuch as scientists use anything like inference to the best explanation, they do not thereby extend our knowledge beyond the observable realm.

Scientific realists, on the other hand, hold that inasmuch as there is a logical gap between explanatory power and truth, this gap can be bridged epistemologically. That is, even if the fact that theory X is the single best explanation for a given class of phenomena does not entail the truth of theory X, the realist claims that its explanatory success nonetheless provides sufficient evidence to justify the belief that theory X is in fact true, or approximately true. Further, the realist will typically attempt to shift the burden of proof by insisting that agnosticism must render the explanatory success of our best theories entirely surprising. Or at a level of meta-abduction, the realist argues that anti-realism makes the overall increasing success of science a miracle, and that therefore realism is the *scientifically* superior philosophical theory. The realist may therefore endorse inference to the best explanation both with respect to individual theories, and in a more general way regarding the overall success of scientific methodology. And he insists that the accompanying epistemic risk is demonstrably worth it.

Before proceeding further, I think it should be stressed that whether scientists actually do make use of these sorts of inferential procedures is not at issue. For almost no one, regardless of how they cast the aims of science, would deny that explanatory inference is a common methodological manoeuvre, or that it has proved itself a profitable tactic. Neither is it controversial that explanatory inference will commonly involve recourse to the unobservable, especially in the modern physical sciences. But for better or for worse, the debate is an essentially philosophical one. It aims at determining the ontological implications of abductive-style inferences by undertaking an analysis of their logical form and epistemological significance. What scientists believe, qua scientists, is beside the point as far as a resolution of the realist - anti-realist conflict is concerned.

This chapter will consist of two sections. In the first section I will explain how epistemological issues have taken center stage in the realist - anti-realist controversy concerning the nature and extent of scientific knowledge. I will show that the problems surrounding scientific realism can (and have been) discussed at the level of metaphysics, semantics and epistemology, but that the most important participants in the contemporary debate are generally in agreement regarding metaphysical and semantic matters. The second part of the chapter will involve an analysis of the inferential strategies used as arguments for scientific realism. I will show that the anti-realist correctly rejects the drawing of ontological conclusions from the methodological procedures of abduction and retroduction. But I will argue further that the deficiencies in abductive and retroductive arguments can be overcome by introduction of the principle of inference to the best explanation.

I.What Is Not At Issue

While the realist - anti-realist debate currently finds its locus in epistemology, this needn't always be the case. There have been, in fact, three general bases of philosophical disagreement between the realist and anti-realist: epistemology, metaphysics and semantics. As I have characterized it, the epistemological dispute concerns the extent to which belief in theoretical entities and processes can be justified. The metaphysical approach turns on questions concerning existence and the nature of truth. From a semantic perspective, settling the issue will depend upon an analysis of the meaning of theoretical and observational terms.

Metaphysics initially enters debates regarding scientific realism with questions about the relation between observation and existence. In general, the more closely one connects these concepts at a metaphysical level, the more sympathetic one will be towards scientific realism. At one extreme, a philosopher committed to Berkeley's dictum *"esse est percipi"* will *a fortiori* be committed to an anti-realist attitude towards theoretical entities. This is also charactersistic of the later phenomenalists, such as Ernst Mach, who reduced the ontology of the world to a complex compendium of sensible elements. Metaphysical positions like these rule out of court the unobservable realm of modern scientific theory.

On the other hand, a metaphysical position which sharply distinguishes the criterion of existence from that of observability will probably, though not necessarily, favour scientific realism. Thus, Platonists and Lockeans will find no metaphysical quarrel with scientific realism, though they may very well find epistemological or semantic grounds for dissent.

A second major point of metaphysical opposition that can arise in the context of scientific realism concerns the nature of truth.[2] For whether or not one thinks the statements of modern scientific theory are true or approximately true will obviously depend on one's conception of truth. For example, "internalists" such as Brian Ellis and (the most recent) Hilary Putnam, who identify truth with "the culmination of the process of investigating and reasoning about nature"[3], must be scientific realists of some form. They are realists by a kind of metaphysical default.

Compare this approach with "incommensurablists" along the lines of Thomas Kuhn and Paul Feyerabend, who contend that the veracity of scientific statements is a theory-relative matter. According to

---

2. It is perhaps not entirely correct to classify 'the nature of truth' as a metaphyscial issue, since truth has traditionally been understood as a semantic notion. My concern, though, is directed at the question of what makes a belief or statement true, and not so much with the nature of truth simpliciter. And an answer to this former question will be influenced by one's metaphysical views. For example, a metaphysical realist will typically claim that statements are true or false by virtue of a correspondence with an independently existing reality. It is in this limited sense that I maintain that metaphysics is relevant to questions concerning the nature of truth.
3. Ellis (1985), p. 69

these philosophers it is not possible "to make a judgement of verisimilitude except from within the confines of a given theory".[4] *Prima facie,* such a view suggests scientific anti-realism since it precludes anything like a trans-theoretic realist conception of scientific knowledge. The anti-realist implications of incommensurability are especially apparent in light of Kuhn's claim that scientists working within different theoretical paradigms actually live in "different worlds".[5]

But metaphysics is not especially relevant to the most prominent participants in the contemporary debate[6], either with respect to the relation between existence and observability, or as regards the nature of truth. The arguments both for and against scientific realism generally assume an unsophisticated metaphysical realism, together with a simple correspondence conception of truth. By metaphysical realism I mean the view that the world exists independently of anyone's beliefs, sensations, or epistemic values. By a correspondence conception of truth I mean the view that a sentence is true or false by virtue of its relation to the world. Given these assumptions, the truth or falsity, referentiality or non-referentiality, of scientific

---

4. Feyerabend (1981a), p. 160
5. Kuhn (1962), p. 111
6. That is, metaphyics is not especially relevant to the arguments of, for example, Boyd, Newton-Smith, and Churchland, on the realist side, or van Fraassen, Laudan, and Cartwright, on the anti-realist side. Metaphysics is relevant to Putnam, Kuhn, and others, but though their insights may have some significance for the realist - anti-realist debate, these philosophers are generally concerned with different issues in the philosophy of science.

theories is not a function of our epistemic faculties, cognitive capacities, or pragmatic desiderata. To this extent, the question of scientific realism, at least as far as I will be concerned, will not be decided upon the basis of any supposed logical or metaphysical relation between existence or truth, and such factors. If they are otherwise relevant to the debate--and we will see that they are--then they will be epistemologically relevant.

I want to make this clear because there is an unfortunate tendency among some commentators to obfuscate the issue by understanding modern scientific anti-realists as holding antiquated metaphysical doctrines. For example, Rom Harre characterizes van Fraassen as a "neo-Berkeleyean" with an even more restricted ontology than the Bishop himself. Harre attributes to van Fraassen the view that "the observable/unobservable distinction is an ontological absolute", and goes on to accuse him of fallaciously drawing ontological conclusions from epistemological premises.[7] But this criticism betrays a naive reading of van Fraassen's brand of scientific anti-realism, for van Fraassen clearly considers the observable/unobservable distinction irrelevant to ontology: "For the term observable classifies putative entities and has logically nothing to do with existence."[8] *Pace* Harre, van Fraassen aims to draw an epistemological, not ontological, conclusion from the observable/unobservable distinction, concerning not what exists simpliciter, but rather what we are justified in believing to exist. His ontology is potentially as rich as anyone's, for he is a metaphysical

7. Harre (1986), p. 58
8. van Fraassen (1980), p. 18

realist: "scientific statements have truth-conditions entirely independent of human activity or knowledge."[9] And scientific anti-realism is compatible with metaphysical realism, just as is any form of skepticism.

I certainly do not want to suggest that metaphysical matters play no part at all in the current debate. Indeed, philosophical theories regarding causality, probability, modalities, universals and the like are in some cases necessary to buttress one's position on the issue of scientific realism. The point is that the central arguments for and against scientific realism do not presuppose a prejudicial metaphysics. There is general agreement among contemporary philosophers of science about what it means to say of some theory that it is true of the world; the fight begins with the question of whether or not we should believe that it is.

Besides metaphysics, the theory of meaning can also influence one's conception of scientific knowledge, in some instances to the extent that semantic issues come to occupy a position of fundamental importance vis-a-vis the realist - anti-realist controversy. I noted earlier that there is a tendency towards scientific anti-realism roughly commensurate with the closeness of connection drawn between observation and existence. In the area of semantics, there is a parallel tendency. In this case, an anti-realist attitude towards scientific theories generally follows from semantic doctrines which closely relate the criteria of meaningfulness to observability

---

9. van Fraassen (1980), p. 38

conditions.

Consider the case of logical positivism, which constituted the dominant philosophy of science for a good part of the twentieth century. The positivists held as one of their central tenets that sentences are meaningful if and only if they are either analytic or contradictory, or their truth or falsity could in principle be determined empirically.[10] With respect to science, this led the positivists to draw a sharp line between observational and theoretical statements, where the former, but not the latter, are understood to admit of confirmation or disconfirmation in a given context through direct observation. This led to the conclusion that the theoretical terms of science are meaningful if and only if they can be translated into language that makes reference only to the observable realm. The consequence of drawing such a close connection between meaning and observation was that reference to unobservable entities had to be construed as a kind of indirect, simplified way of referring to observables in order to qualify as significant. So according to positivist semantics, reference to unobservables must either be understood in a non-literal sense, or dismissed as meaningless. Either way, scientific realism loses out.

I have argued that positivist theories of meaning, which connect closely the criteria of meaning and observation, will require a sharp observation-theory distinction, and consequently lead to a form of scientific anti-realism. I think the opposite is true of philosophers who, in reaction to positivism, play down the significance of the

---

10. Hempel (1965), p. 101

distinction.[11] Wifred Sellars, for example, attempts to undermine the 'received view' of the conceptual structure of scientific theories, according to which the observation framework garners an absolute status independent of thought and language. Sellars argues that, for a given theory, the observation framework in which its empirical consequences are couched will not be wholly free of theoretical contamination. Whatever inductive generalizations can be sustained at the observation level as empirical consequences of the theory can be sustained "only by a covert introduction of the framework of the theory into the observation framework itself."[12] Therefore, theoretical considerations are tied into the meaning of observation statements, and indispensable to the functioning of the theory as an explanatory device. Hence there is no conceptual or semantic justification for drawing an ontological line at the boundaries of the observation framework. On Sellars' view there can be no reason for dismissing the ontological implications of a theory at the unobservable level that is not also a reason for rejecting its empirical consequences at the observable level, and vice-versa. This, I think, is what is behind Sellars' well-known realist manifesto that "to have

---

11. The earliest and most comprehensive post-positivist attacks on the observation-theory dichotomy were those of N.R. Hanson (1958), and Peter Achenstein (1968). Sellars' brought the significance of the dissolution of this dichotomy to bear on the philosophy of mind and epistemology. In this regard, see his articles, "Empiricism and the Philosophy of Mind" (in Sellars (1962)), and "Scientific Realism or Irenic Instrumentalism?" (Sellars (1965)).
12. Sellars (1976), p. 315

good reason for holding a theory is *ipso facto* to have good reason for holding that the entities postulated by that theory exist."[13]

The positivist and Sellarsian views concerning the semantic and conceptual status of theoretical and observational terms effectively rule out *ab initio* the possibility, respectively, of a realist or anti-realist attitude towards scientific knowledge. The positivists' emphasis on direct observation requires a non-literal interpretation of theoretical terms. On the other hand, Sellars' insistence on the theory-infectedness of the observation framework makes questions about, say, the existence of atoms identical to questions about the explanatory adequacy of atomic theory as a whole. Neither approach will allow for anything like the moderate realism or limited skepticism that characterize contemporary positions in the realist - anti-realist controversy.

But just as we have found that metaphysical issues do not occupy a central place in the current debate, neither do semantic ones. The most prominent participants on either side assume what might be called semantic realism regarding scientific statements. On this view, both observation terms and theoretical terms have the capacity for genuine reference. Thus, Larry Laudan prefaces his argument against epistemological realism by stressing that there is no dispute that theories have determinate truth-values.[14] Similarly, van Fraassen's form of scientific anti-realism presupposes a literal understanding of scientific language whereby "the apparent statements of science

13. Sellars (1962), p. 97(n)
14. Laudan (1984), p. 219

really are statements, capable of being true or false".[15] Contemporary anti-realists for the most part eschew positivist semantics; their arguments are primarily epistemological.

Nor does much turn on the so-called theory-infectedness of observation noted by Sellars and others. It is generally agreed that the inability to draw a very sharp semantic-conceptual line between theoretical and observational statements leaves open the question of whether or not an important epistemological line can be drawn. Thus, van Fraassen agrees that all of our language is theory-laden, but contends that one can, in an admittedly anthropocentric fashion, nonetheless "classify objects and events into observable and unobservable ones."[16] The scientific realist Richard Boyd agrees that the failure of the positivists to construct a theoretically neutral sense-datum language does not entail that the line between the observable and the unobservable is not epistemologically significant.[17]

Members of both camps seem to share the opinion of the entity-realist Ian Hacking that "whatever be the interest in the philosophy of language, it has little to contribute to our understanding of science."[18] Indeed in his recent book on scientific explanation we find the realist Salmon explicitly endorsing van Fraassen's declaration that "the main lesson of twentieth-century philosophy of science may well be this: no concept which is essentially language dependent has

15. Van Fraassen (1980), p. 10
16. Ibid p. 14
17. Boyd (1984), p. 45
18. Hacking (1983), p. 45

any philosophical importance whatsoever."[19]

It should be clear from the preceding discussion that the realist - anti-realist battle can be fought on three different fronts. This is not to suggest that any aspect of the debate is primary or fundamental. In fact, a complete theory of scientific knowledge will require victory on all fronts. Thus, a comprehensive scientific realism must show not only that belief in theoretical entities is epistemologically justified, but also that phenomenalist metaphysics and the positivist theory of meaning are untenable. I think it should also be kept in mind that the metaphysical, semantic, and epistemological analyses are not entirely independent of one another, and that a position on one aspect of the debate may suggest, or even entail, a position on another aspect. For example, idealism entails the falsity of any epistemological defence of scientific realism, while incommensurablism suggests, but does not entail, anti-realism.

But I do not intend to provide a complete philosophical analysis of the nature of scientific knowledge. I have so far argued only that the epistemological debate, concerning the justification for belief in unobservables, need not presuppose controversial metaphysical or semantic doctrines. It is therefore neither metaphysics nor the theory of meaning that are at issue as far as the contemporary debate, and this thesis, are concerned. As for epistemology, the central question I shall deal with is this: to what extent does the explanatory power of scientific theories warrant belief in the entities and processes postulated by those theories?

19. Salmon (1984), p. 91(n)

II. What Is At Issue

Larry Laudan has remarked that the fundamental motivation for scientific anti-realism is "the deep-rooted conviction that the fallacy of affirming the consequent is indeed fallacious."[20] Laudan is here referring to the logical structure of abductive inference. I earlier noted that abductive arguments contain two premises. The first premise describes some observed state of affairs (B). The second states how some other state of affairs (A) might account for the phenomena: (A->B). The conclusion drawn is that (A) holds. Laudan is of course correct that the argument is deductively invalid. But this is no surprise--abduction isn't deduction, and should no more be expected to conform to the rules of deductive inference than should induction. The question is, can abductive arguments be justified even if they do not provide strict entailment? Consider an everyday sort of example. Returning home one day, you notice that the lights of your apartment are on. 'My roommate is home early', you decide. The reasoning involved is abductive since it proceeds from the knowledge that if she were home the lights would be on, to the conclusion that she is in fact home. This a clear case of affirming the consequent, but appears to be justified nonetheless. It therefore cannot be simply the deductive invalidity of abductive reasoning that motivates the anti-realist's agnostic stance. It would be a heroic form of skepticism that restricted belief to only those conclusions that are arrived at deductively.

---

20. Laudan, (1984), p. 242

Still, I think Laudan has located the heart of the matter regarding the proper epistemic attitude to hold towards the unobservable postulates of theoretical science. Consider again the roommate example. Our confidence in the conclusion is based, I think, on the assumption of certain prior probabilities. The alternative explanations for the lights being on (thief in the house, electrical malfunction, I forgot to turn them off, etc.) are assigned a relatively low antecedent probability given our background knowledge (the safety of the neighbourhood, the soundness of the apartment's wiring system, etc.). Our conclusion is therefore justified relative to the comparatively small likelihood of alternative explanations.

But in theoretical science hypotheses are originally devised specifically to account for the given phenomena. There is therefore no background knowledge concerning the behaviour of the entity or process postulated that might suggest an antecedent probability that they are operative in the present context.[21] Peirce writes that the

---

21. It may be objected that this is an extreme idealization. In actual scientific contexts, a great deal of related background theoretical knowledge is assumed, which would tend to suggest prior probabilities concerning what sort of entities are operative in the given experimental context. Furthermore, technological advances in fields such as electron microscopy have made possible direct empirical verification of some explanatory hypotheses. Therefore, the differences to which I refer between our everyday use of abduction, and those of the scientist, are not so significant after all.

The problem with these objections is that they may beg the question in assuming that the information provided by background theories counts as genuine knowledge, and that we really do see through an electron microscope. Nonetheless, these points certainly abductive inference should not be based on any previous knowledge that might bear on the truth of the hypothesis.[22] Nor, if they invoke unobservables, will the candidate hypotheses be subject to verification through direct empirical means. The only thing, if anything, that will be relevant to a judgement concerning the truth of any single candidate hypothesis, is the question of whether or not the hypothesis, if true, would produce the observed effect. And since any number of hypotheses will meet this criterion, and in lieu of a prior probability ordering, inference to the truth of any one cannot be justified. For example, in the absence of some independent antecedent ground for supposing that one of either a wave or a particle is operative in a given situation, and to the extent to which the presence of either would be expected to give rise to the observed phenomena, the conclusion that one as opposed to the other is at work will not be justifiable. Abduction alone lacks the epistemic power to warrant ontological conclusions, at least so far as theoretical science is concerned. This I take to be the point of Laudan's objection to the consequent-affirming character of the abductive inference.

Nonetheless, in scientific contexts abduction is meant to serve only as a means identifying a range of candidate explanatory hypotheses, and by no means exhausts the resources of scientific methodology. A second, distinct process which, again following Peirce, I shall call retroduction, is required before acceptance of any one of the candidate hypotheses can be warranted. Retroduction

---

do pull in the direction of realism, and are, in fact, the basis of Hacking's excellent 'argument from the grid'. ((1983), ch. 11)

22. Peirce (1934), 6:526

begins with the candidate hypotheses arrived at abductively, and subjects them to experimental testing. The general strategy is, for each hypothesis, to make some prediction based on the assumption that the hypothesis is true. An experiment is then conducted to determine whether or not the prediction is born out. If it is not, then the hypothesis is considered falsified (by the formal argument of *modus tollens*) and discarded. If the prediction is confirmed, further tests are undertaken. By this process of elimination, the scientist hopesfinally to arrive at a single hypothesis that is fully consistent with the experimental data.

Although retroduction enables the scientist to falsify candidate hypotheses by showing through experiment that their observable consequences do not obtain in all relevant cases, and is to this extent superior to simple abduction, it still cannot verify hypotheses. For it does not follow from the fact that an hypothesis has only true consequences that it is true itself. To say that it did follow would be to affirm the consequent. I think therefore that we should understand abduction and retroduction as methodological principles, reliable towards the end of providing the best theoretical explanations, but unable to guarantee the truth of any particular theory. So again the question arises as to whether or not we can be justified in believing the best theories to be true, or nearly true, even if their truth is not entailed by their being the best theory.

The realist argues that an affirmative answer can be justified via the principle of inference to the best explanation. We have so far seen that abduction and retroduction will lead to the acceptance of the best scientific hypotheses, but do not themselves warrant conclusions regarding their truth or falsity. Abduction identifies a finite range of explanatory hypotheses, but does not select from among them. Retroduction isolates the single best explanatory hypothesis by a process of falsification. We thus arrive at a theory which is first of all explanatory, and second, consistent with all of the data. Yet so far no ontological conclusions can be drawn beyond the level of the observable since the falsity of a theory is compatible with its surviving this dual procedure. However, if it can be argued that the explanatory success of the surviving theory is itself best understood by assuming it to be true, or approximately true, then the realist may finally be justified in drawing her ontological conclusions. This is the general strategy behind the principle of inference to the best explanation.[23]

If tenable, the principle of inference to the best explanation will fill the epistemic gap between explanatory success and truth left by abduction and retroduction. Consider the premises of the abductive argument form: (1) B (the phenomenon in need of explanation)

---

23. There is a certain amount of ambiguity in the name of this principle. For as Ernan McMullin has pointed out, ((1978), pp. 222-223), the inferential direction is from the claim that a certain hypothesis is the best explanation, to the conclusion that the hypothesis is true. For this reason, he suggests that the principle would be more aptly named inference *from* the best explanation. On the other hand, it may be said that there is no logical difference between inferring that the best explanation is true, and simply inferring to the best explanation. The ideal name for the principle is probably 'the inference from the claim that a hypothesis is the best explanation, to the truth of the hypothesis'. This noted, I will continue with the standard terminology.

(2) A -> B (hypothesis A will explain the phenomenon B)

We now add a third premise from the results of retroduction:

(3) Ab -> B (only hypothesis A gives rise to no false consequences in experimentation)

From these premises we may conclude that A is the single best explanatory hypothesis available, and perhaps may also be entitled to the methodological prescription that it ought to be accepted.[24] But to conclude further that A is true will require a fourth premise stating that A's being the best explanatory hypothesis justifies the belief that it is true. To this end, inference to the best explanation combines the findings of the methodological procedures of abduction and retroduction with the epistemological assumption that explanatory success warrants belief, to arrive at the realist conclusion that our best scientific theories are not only methodologically reliable, but also true, or approximately true.

What could justify a fourth premise recommending inference to the best explanation? There is an argument implicit in the inference to the best explanation that can be formulated in the following way:

(1) Ab (theory A provides the best explanation of the given phenomenon)

(2) At -> Ab (true theories typically will provide

24. There may, of course, be other factors which enter into the question of whether a certain hypothesis should be accepted. For example, is it simpler than available alternatives? For now, though, I am interested only in the method by which the best hypothesis, qua explanation, is determined.

better explanations than false theories)

(3) Ab -> At (theories that provide the best explanations typically will be true)

(4) .'. At (A is probably true)

(5) .'. Abe (A ought to be believed)

The first premise re-states the conclusion of the procedures of abduction and retroduction. The second premise simply asserts what is anyway assumed by retroduction, that a true theory will have only true consequences. (We will see that some anti-realists find issue with this sort of premise). (4) follows from those that precede it, and (5) follows from (4) by a common-sense epistemological inference. For most anti-realists, the contentious premise is (3), which, together with (2), formally establishes a biconditional between truth and explanatory success. The significance of this is that it solves the problem of affirming the consequent that plagued attempts to draw ontological conclusions from the premises of abductive and retroductive arguments relating the explanatory power of some theory or hypothesis. For if it can be assumed that the best explanatory hypotheses will typically be true or nearly true, then once we have arrived at the best theory through the usual methodological procedures, we will be justified not only in accepting it, but in believing it as well. When I come to discuss the main anti-realist objections to inference to the best explanation, we will see that this is primarily where the attacks are directed.

There are two general levels upon which inference to the best explanation can proceed as part of the realist stategy of establishing an epistemological connection between truth and explanatory power.

On the basic level, inference to the best explanation is theory-specific, prescribing belief in any theory within a mature science that has shown itself to provide the best overall explanation for the phenomena within its domain. We will see that there are good reasons for the realist to avoid this approach.

On a meta-level, inference to the best explanation is invoked towards showing that the overall increasing explanatory success of science justifies the belief that science, as a whole, is approaching an ever more accurate description of the world at both the observable and unobservable levels. This is the strategy adopted by most prominent contemporary scientific realists like Boyd, Putnam, Devitt, and McMullin. There is also usually a rhetorical side to this approach whereby the realist maintains that her view is the only philosophy of science that does not render the success of science a miracle. Such realists therefore like to characterize their position as an over-arching empirical hypothesis, a plausible explanation for the success of science to which the anti-realist can offer no alternative.

Though these strategies differ in focus, they share at bottom the fundamental realist conviction that explanatory success is a reliable indicator of truth and referentiality. The anti-realist aims through various avenues at revealing the untenability of this conviction, and at showing that nothing but some bad philosophy is lost in divorcing the search for explanation (and hence science itself) from the search for truth. In what follows, I will evaluate the realist - anti-realist debate as regards this fundamental point of opposition.

CHAPTER TWO:
EXPLANATORY INFERENCE: OBJECTIONS

Introduction

In this chapter I discuss the two most important arguments that have been offered against the scientific realist's employment of the principle of inference to the best explanation. The first argument is essentially logical, taking the form of *reductio ad absurdum,* and purportedly demonstrating that adherence to the principle of inference to the best explanation will lead to the recommendation of belief in the truth of mutually incompatible theories. This is the argument from the thesis of the empirical underdetermination of theory by data. The second major argument is empirical, and based on a meta-induction from the history of science. It attempts inductively to undermine inference to the best explanation through a sampling from the scrap-heap of past scientific theories which once provided the best explanation for the given domain of phenomena. Since there are no general epistemological grounds for believing in the truth of contemporary scientific theory that were not equally good reason for believing in the truth of past theories, and since we now know that most or all of those past theories were radically off the mark, it would require something akin to a scientific leap of faith to believe in the truth of our present theories. That is, the lessons of history constitute sufficient inductive support for the view that even our best scientific theories will turn out to be false, and hence that inference to the best explanation is unjustified.

I aim to show that together these arguments succeed in defeating the strategy of employing inference to the best explanation towards any theory-specific form of scientific realism. But I will contend also that the more general inferential strategy is left relatively unscathed. This is the stategy, charcteristic of the 'no miracles' argument, of setting scientific realism itself as the best explanation for the overall increasing success of science.

I. The Empirical Underdetermination Thesis

It was shown in the first chapter that, by themselves, the dual methodological principles of abduction (hypotheses formulation) and retroduction (hypothesis selection by experimental falsification), will not justify belief in the unobservable aspects of scientific theory. While they are sufficient to identify a theory that provides a consistent explanation of the relevant phenomena, and therefore to recommend its acceptance by the scientific community, they do not warrant the further belief that the theory is true. For, quite simply, nothing concerning the justifiability of this belief follows from the knowledge that a certain theory constitutes the best available explanation of the given phenomena. To justify this further belief requires an epistemological principle that somehow connects truth and explanatory power. This is supplied by appeal to the principle of inference to the best explanation, which prescribes belief in the theory that survives the methodological procedures, on the assumption that explanatory success serves as a reliable indicator of truth.

Without even entering into the debate about whether or not explanatory success is a reliable indicator of truth or reference, there is an important and extremely elegant argument which suggests that adherence to the realist's inferential principle will lead to unhappy logical consequences. The realist maintains that belief is justified primarily by the fact that some theory has only true empirical consequences relative to the known data within its theoretical domain, i.e., that it provides the best explanation. But for any such theory, there may exist a second theory which is empirically equivalent to the first (has all and only the empirical consequences of the first) but which is distinct and ontologically incompatible at the unobservable level. Given this, any reason that may be held as reason for belief in the truth of the first theory will count equally as reason for belief in the second. Therefore, it would seem that the principle of inference to the best explanation must counsel belief in the truth of incompatible theories. *Reductio Ad Absurdum.* This is known as the argument from the empirical underdetermination of theory by data.

This argument has a long and popular history in the tradition of scientific anti-realism, from the early positivists down to contemporary constructive empiricists. It was used by Duhem, Poincare and Mach, and is the central argument invoked against realism by van Fraassen in The Scientific Image. Duhem, for example, in a discussion of the significance of experimentation, writes that the process of empirical testing,

> does not have the power to transform a physical hypothesis into an undisputable truth; in order to

> confer this power on it, it would be necessary to enumerate completely the various hypotheses which may cover a determinate group of phenomena. But the physicist is never sure he has exhausted all the imaginable assumptions. The truth of physical theory is not decided by heads or tails.[1]

It was reasons such as these that led Duhem to the conclusion that the proper aim of science was the ordering and classification of phenomena, and that explanation, and the determination of the reality that lay beneath the appearances, was the role of metaphysics. Poincare argued similarly that the relations scientists postulate to hold between two bodies A and B "could just as well have replaced advantageously a relation that holds between to other bodies A" and B", entirely different from A and B."[2] Considerations such as this supported Poincare's conventionalist thesis that the sole object of scientific inquiry was the coordination of phenomenal laws arrived at through observation and experiment.

Van Fraassen makes similar use of the underdetermination problem to buttress his claim that the proper aim of science is not truth but empirical adequacy, deriving cases of empirical equivalence from the theoretical domains of both classical and quantum mechanics. He argues, for example, that Newton's general theoretical framework will have all of the same empirical consequences regardless of whether the postulate of Absolute Space is assumed to be at rest (TN-0), as Newton in fact thought, or at some

---

1. Duhem (1954), p. 190
2. Poincare (1958), p. 126

constant velocity (TN-V). Though the postulation of Absolute Space was essential to Newton's account of the relative motions of celestial bodies, the question of whether or not Absolute Space is itself in motion is underdetermined by whatever empirical data can be expected to follow from the statements of the theory. Therefore, van Fraassen maintains, it is rational to accept each version of the theory as empirically adequate. But it is not rational to believe either given that there can be no empirical grounds for the truth of one, as opposed to the other, and given that believing them both is contradictory.[3]

Consider also an argument of George Schlesinger, which he attributes to H. Jeffreys, concerning the path of planetary orbits. Schlesinger points out that, for any finite amount of data concerning the position of a planet along the path of its closed orbit, it is always possible to construct an infinite number of different geometric models that will fit the given information. And this will be true no matter how much data is amassed. Schlesinger's solution involves appeal to a principle of maximizing simplicity. I will argue later that such a move does not really help matters as far as realism is concerned. At any rate, in the face of such radical underdetermination, agnosticism would seem, *prima facie*, to be the only rationally defensible epistemic position.[4]

As an argument for scientific anti-realism, the thesis of emprical underdetermination derives its strength from the fact that it attacks

---

3. van Fraassen (1980), pp. 44-46
4. Schlesinger (1974), p. 34

directly the inferential move adopted by realists from explanatory power to truth. While it seems obvious that theories determine (in the sense of predict) certain specific data, it is perhaps equally obvious that data do not determine any particular theory. There is nothing logically incoherent in the idea that more than one theory may provide a complete and consistent explanation for a given empirical subject matter. And if it is possible that all of our best scientific theories are in this way radically underdetermined by a potentially infinite number of alternative explanations, abductive inference to the truth of the theories actually currently held begins to look unreliable, if not altogether arbitrary.

But the logical point that scientific theory is not determined by data leaves open the more important question of the epistemological significance of this underdetermination regarding the realist - anti-realist debate. I say this question is more important because there are several ways in which the underdetermination of theory by data can be a trivial matter. Before assessing the impact of underdetermination difficulties on scientific realism, I first want to note and dispose of these harmless forms of underdetermination.

First, there is a sense in which even our most basic, pre-scientific, beliefs about the world are underdetermined. For example, my belief that I am now in Calgary on a winter evening has as one empirical consequence that I will feel chilled if I go outside without a sweater (chinooks notwithstanding). But the same consequence will follow if I am under the control of a deceiving god, or actually in Hawaii dreaming that I am in Calgary. In fact, skeptical hypotheses such as these are designed to be empirically equivalent to, but incompatible with, all of our common sense beliefs about the world. They are intended to undermine our confidence in such beliefs by showing them to be underdetermined with respect to all possible observation.

This sort of wholesale underdetermination is an immediate consequence of any metaphysical realist position, and this fact has led some philosophers to abandon that doctrine.[5] But, as I argued in the first chapter, the contemporary brand of scientific anti-realist is avowedly realist in her metaphysics. The sort of underdetermination suggested by radical skepticism is therefore not at issue in the context of the present debate. For scientific anti-realists admit a far wider range of 'safe' empirical data than does the skeptic, for whom only immediate appearances and the like are not up for doubt. This radical form of underdetermination, though certainly interesting in its own right, is therefore peripheral to the realist - anti-realist debate concerning the extent of scientific knowledge.[6]

---

5. Ellis (1985), pp. 66-67
6. I do not mean to suggest that wholesale underdetermination has no importance at all in the philosophy of science. We should remember that the arguments of van Fraassen and his allies are essentially skeptical. But the anti-realist about theoretical knowledge, especially if she has empiricist leanings, wants to restrict her skepticism to the unobservable realm. Still, it may be argued that there is no philosophically significant difference between the empirical underdetermination of theoretical entities, and that of ordinary observable objects, and therefore that a thoroughgoing anti-realism must end either in wholesale skepticism, or some version of internal realism such as that adopted by Ellis and Putnam. This is exactly the line of argument taken by some realists against van Fraassen (Churchland (1985), Gutting (1985), Musgrave (1985)). For my purposes, I will admit that there is an important epistemological

A second form of underdetermination that is harmless to scientific realism follows from a broad interpretation of the condition of empirical equivalence. For two theories which are distinct and incompatible at the unobservable level, and which have no empirical content whatsoever--generate no empirical predictions even in conjunction with other auxiliary conditions--will be empirically equivalent, and therefore underdetermined. Consider two theories, T' and T". T' says simply that the Ego is transcendent, while T" says that the Ego is immanent. Though incompatible ontologically, these theories are empirically equivalent for the simple reason that neither has any empirical consequences at all. It may be argued that most moral and metaphysical theories are underdetermined in this sense. We may legitimately require that scientific theories, to count as scientific, have some empirical content, and so conclude that this form of underdetermination in no way threatens scientific realism.

There is a related way of constructing empirically equivalent theories that is similarly harmless. Take some theory T with the set of empirical consequences (E1. . . En). Now devise two further theories, T' and T", by adjoining to T in each instance further hypotheses that are mutually incompatible but empirically empty. T, T', and T", by virtue of the addition of the gratuitous, empirically empty, hypotheses, are all empirically equivalent, and hence underdetermined by the relevant data. By this method, we can construct empirically equivalent counterparts for any scientific

---

distinction to be drawn between the observable and the unobservable, and concentrate on the question of whether or not it can be bridged by principles of explanatory inference.

theory. For example, the conjunction of quantum mechanics and some moral theory P will have all of the same consequences as quantum mechanics alone. It should be clear that this frivolous addition of empty hypotheses does not affect the question of the nature of scientific knowledge, if for no other reason than that scientists generally do not engage in constructing theories by this sort of gratuitous addition to the statements of existing theories.

Finally, it has been illustrated by Quine[7] that empirically equivalent but logically incompatible theories can be constructed simply by undertaking a systematic switching of two central theoretical terms. Thus, we might replace "neutron" with "proton", and vice-versa, wherever they occur in some standard formulation of electron physics. The 'new' theory will retain all of the same observational consequences as the original, but will be, at least logically, distinct and incompatible: it will confirm things about so-called protons that the other will not, for example. Quine argues that despite the strict logical incompatibility, the two should nonetheless be regarded as the same theory, since the two formulations can be rendered logically equivalent by a simple reconstrual of the terms "proton" and "neutron", without affecting the meaning of any theoretical statement in which they are couched. This sort of underdetermination is therefore no problem, since the two theories that are supposedly underdetermined should in fact properly be understood as logically distinct formulations of the same theory. The significance here is that mere logical incompatibility is not sufficient

7. Quine (1975) p. 319

for a genuine case of empirical underdetermination. It is also necessary that the theories be ontologically distinct.

From the preceding survey of trivial forms of empirical equivalence, we can devise a list of conditions that must be met by a putative case of empirical underdetermination for it to pose a serious threat to scientific realism:

1. The theories must have all and only the same empirical consequences.
2. The theories must be ontologically incompatible at the unobservable level.
3. The theories must not underdetermine beliefs concerning the observable realm.
4. The theories must have some empirical content.
5. The theories must not include any gratuitous, empirically empty hypotheses.

Since we have not yet ruled out the possibility that there are other criteria besides empirical success that can stand as evidence for the truth of a theory, we should add the further condition that:

6. The theories must be *evidentially* equivalent or indistinguishable.

There are at least two ways by which philosophers have attempted to *dissolve* the problem of theory underdetermination. First it has been argued by some that empirically equivalent theories are synonymous, and therefore contradict one another at the unobservable level only in appearance, but not in fact. This way of getting around the problem derives from a positivistic interpretation of the meaning of theoretical terms. On this view, the factual content of a theory is exhausted by its empirical import, and theoretical terms are judged meaningful if and only if they can be translated without remainder into observation statements. Therefore, inasmuch as two theories are empirically equivalent, they will also be factually equivalent and have the same truth values. The choice between them will be a matter of convention. Hans Reichenbach illustrates this approach with the example of measurement systems:

> There are cases in which the theories compared are logically equivalent, i.e., correspond in all observable facts. A well known case of this type is the difference of the systems of measurement. The metrical system is simpler than the system of yards & inches, but there is no difference of truth character.[8]

The choice between two such systems is for Reichenbach a practical matter, and he suggests that the choice between empirically equivalent theories in modern physics is likewise irrelevant to considerations of truth and falsity.

There are three reasons why this way of dissolving the problem of underdetermination is unsatisfactory. First, no attempt at a translation of theoretical terminology into a purely observational vocabulary has been successful. Secondly, scientific realists think the truth of theoretical statements to be semantically independent of the truth of observation statements. They think that theoretical truth is indicated by, but not defined by, empirical facticity. This positivistic way of avoiding the problem of underdetermination can therefore

---

8. Reichenbach (1938), p. 374

offer only cold comfort to a scientific realist who intends his thesis in a robust sense. Finally, as I argued in the first chapter, contemporary scientific anti-realists agree that theoretical terms should be construed literally, and would as a result deny that this sort of move saves the brand of scientific realism they oppose. We should therefore reject the synonymy solution to the problem of empirical underdetermination.

A second argument aimed at dissolving the problem of underdetermination goes as follows. The problem concludes that in situations where two or more theories agree entirely with respect to observable claims, but conflict in their theoretical claims, there can be no reason to believe in the truth of one of the theories that is not equally reason for believing in the truth of the others. But this presupposes a difference *in kind* between theoretical and observational propositions or judgements. And since all observation is theory-laden, no theory-observation dichotomy is legitimate. Therefore the problem cannot even be formulated.

The argument goes too far. It is certainly true that all observation is theory-laden in the sense that no empirical judgements are free of conceptual classification. Denying this would lead to an indefensible form of foundationalism. To this extent, all of our beliefs, including our beliefs about observables, are theoretical. But this is no problem for the scientific anti-realist, for she draws the line of justified belief at the boundary of the observable, not the theoretical. And even if there is no sharp line between the observable and the unobservable, there are certainly paradigm instances of each. Thus, 'microwave oven' and 'electron' are both theory-informed concepts, though they clearly respectively denote observable and unobservable entities.[9] And while we may disagree, it clearly would be wrong to think that a person's decision to believe in the former, but not the latter, to be simply arbitrary. There seems to be an important epistemological difference between observables and unobservables, even if there is no difference *in kind* between theoretical and observational concepts. Therefore, the anti-realist need not presuppose that observation is not theory-laden to argue that the unobservable aspects of scientific theory are underdetermined by their observable consequences, and hence that belief in the former cannot be justified.

Are there any cases of empirical underdetermination in modern science that meet the conditions set out above? Quantum mechanics, with its competing interpretations of the nature of the processes underlying certain systems, seems to be a good candidate. And there are surely others as well. But just as surely, there are many theories that are not so obviously underdetermined, such as the genetic theory of the transmission of human traits, or the kinetic theory of temperature. If the debate about underdetermination was conducted

---

9. This is not to say, of course, that when we observe a microwave oven, we observe microwaves. Rather, one of the properties that constitute our concept of a microwave oven, 'emits microwaves', is unobservable. It also includes many observable properties. On the other hand, our concept of an electron includes no observable properties. The point is that the theory-ladeness of observation does not entail that there are not observable and unobservable properties which objects can possess. The anti-realist denies that we can have genuine knowledge of objects which have no observable properties, which happens to be the case with most of the postulates of the physcial sciences.

simply at the empirical level, it is not clear that it would have any significance for scientific realism. For when presented with putative cases of underdetermination, the response is always available on behalf of realism that all the data may not be in, or that future theoretical developments may yet favour the truth of one of the underdetermined theories. The realist may thus continue to adhere to the principle of inference to the best explanation, while at the same time admitting that in some cases it is not *now* clear what the best explanation is.

The anti-realist must therefore intend the thesis of empirical underdetermination not as an isolated fact about some theories presently held, but rather as a logical claim to the effect that for any given theory at any given time it may (for all we know) be underdetermined at the unobservable level. The point of citing actual examples is to show that underdetermination, being actual in some cases, may be possible in all cases. Understood in this sense, the realist response will have to suggest the following: for any given theory such that there may exist another theory that is empirically equivalent but ontologically incompatible, there is some potential means of distinguishing between them evidentially, thereby justifying belief in one over the other. I now want to consider two basic ways in which realists have attempted to show how empirically equivalent theories might nonetheless be differentiated on other grounds relevant to judgements about truth and falsity.

The first argument focuses on the role of auxiliary hypotheses in experimentation. Theories are rarely, if ever, tested in insolation. Typically, predictions are derived on the basis of the theory in conjunction with certain accepted auxiliary hypotheses. Thus, two theories may be empirically equivalent--entail all of the same observational consequences--yet yield different predictions when considered together with certain other theoretical assumptions. For example, two empirically equivalent theories may postulate, respectively, wave and particle functions at the unobservable level. Now imagine some third well-confirmed theory which says, among other things, that the sort of entity postulated by the second theory undergoes a radical change in behaviour above a certain temperature. Together with this auxillary hypothesis an experiment may be designed from which we can expect different data, depending on whether there is, in fact, a wave or a particle operative in the given situation. Therefore, there will be empirical evidence that will count in favour of the truth of one of even empirically equivalent theories.

This objection has the virtue of reflecting the actual character of scientific research. For without including a host of background assumptions, some of which will be theoretical, the abstract isolated formulation of a theory will have few non-trivial empirical consequences. But though the objection serves to shed light on an important aspect of scientific methodology, it manages only to push back one step the argument of the anti-realist regarding empirical underdetermination. While it may be that empirically equivalent theories will yield different experimental predictions when they are conjoined with with certain accepted auxillary hypotheses, the claim that this provides evidence for the truth of one of the theories assumes just what is at issue: that some theories can be justifiably believed to be true. Therefore, the force of the realist's appeal to the role of auxillary hypotheses depends already on the falsity of the universal underdetermination of theory by data, And, as I have argued, this is how the anti-realist must intend the thesis. This attempt to show how empirically equivalent theories might still be evidentially distinguished manages only to the beg the question against the anti-realist.

A similar response is that empirically equivalent theories may be distinguished according to which of them does, or does not, conflict with a more general theory in the relevant area of research. Thus it is urged that of two equivalent theories, T' and T", which postulate radically distinct kinds of forces, that which does not contradict a more general theory of forces F, is more likely to be true. Again, this looks likes a sound methodological prescription, but fails to address the anti-realist's claim that *all* theories, and hence all research traditions, may be underdetermined by undiscovered or unformulated alternatives. The realist's claim that the successful extension of one of the empirically equivalent theories to related or more general systems counts as grounds for believing it more likely to be true, assumes precisely what the anti-realist denies--that we can have justified belief regarding the unobservable aspects of some scientific theory.

Finally, some realists have tried to preserve inference to the best explanation in the face of the underdetermination problem by broadening the criteria for 'the best explanation' beyond mere empirical support. Thus it is argued that the anti-realist ignores the so-called 'super-emprical virtues' which may be enlisted towards evidentially distinguishing between empirically equivalent theories. The strategy is to identify one of two or more empirically equivalent theories as *prima facie* more likely to be true given either the present theoretical structure of science, or some generally accepted metaphysical schema. It may be argued, for example, that of two empirically equivalent theories T' and T'', it is more likely that T' is true because, unlike T'', it fits nicely into the dominant theoretical framework of contemporary science, or because it is far simpler than T''.

It should be noted that the realist's contention here is, and must be, stronger than merely that the super-empirical virtues provide a legitimate means for choosing between empirically equivalent theories. For the anti-realist will certainly admit that such considerations are central to the methodology of theory-choice. What the anti-realist denies is that they count as anything more than pragmatic grounds for choosing between empirically equivalent theories. Therefore, in order for the appeal to super-empirical virtues to count as a serious rebuttal to the anti-realist argument from underdetermination, it must be intended as supplying reasons for *believing*, as well as for choosing, one of the underdetermined theories. The realist must think that such factors as relative simplicity, and agreement with the established body of theory, are relevant to judgements concerning truth or falsity.

But it is difficult to see how such a reply to the underdetermination argument can get off of the ground without once again begging the question against the anti-realist. For it is not clear how the epistemic relevance of the non-empirical considerations can be supported in a non-dogmatic way. Take the supposed virtue of simplicity. It is a Humean kind of point that the assumption of the fundamental simplicity of nature must either be based on experience, or on metaphysical dogma. If it is the former, then it will be a matter for scientific investigation, and hence can not be adduced as a reason for thinking a theory to be true or false over and above the usual empirical criteria of theory choice. And since, by hypothesis, empirically equivalent theories will be on a par in this regard, whether or not nature is simple in fact will depend on whether or not the simpler theory is true. And of course the question of whether or not this can be determined is precisely what is at issue. On the other hand, if the assumption of simplicity is a purely dogmatic metaphysical principle, it is not defensible against contrary dogma, and will not therefore be relevant to the question of which of two empirically equivalent theories is more likely to be true.

The same kind of dilemma emerges for the realist who suggests that agreement with the accepted theoretical framework counts as a non-empirical ground for believing a theory more likely to be true than another which contradicts the theoretical status quo. If this criterion is to be supported empirically, it will depend on the assumption that belief in the truth of the more general theoretical framework with which the theory accords can be justified empirically. But the anti-realist insists that the underdetermination argument shows this sort of assumption to be unjustified. And so a regress ensues. And the regress can be stopped only by dogmatically asserting the truth of some theory, so that of two empirically equivalent theories, the one which accords with *it* can be presumed to be more likely to be true. We should conclude that the realist's appeal to the super-empirical virtues must depend either on a dogma prejudicial against anti-realism, or finally on empirical factors. In either case, the argument from empirical underdetermination is not seriously affected.

I conclude that the various realist responses are unsuccessful against the logical point of the underdetermination argument, which is that for any given theory at a given time that enjoys substantial empirical support, there may exist some as yet undiscovered or unformulated alternative with which is ontologically incompatible, but from which it is evidentially indistinguishable. This alone does not show the falsity of scientific realism, nor does it even undermine the general appeal of inference to the best explanation. For, as I will argue in the third section of this chapter, there is a meta-version of the inference to the best explanation that is not subject to the problem of empirical underdetermination. What we must conclude is that any attempt to infer from the fact that some theory gives the best available explanation for the empirical data to the truth of that theory, is rendered epistemologically suspect by the argument from underdetermination.

II. The Historical Gambit

A second important challenge to the realist's employment of the principle of inference to the best explanation is based on an induction from the unhappy history of scientific theory. The realist strategy of connecting explanatory success and truth is seriously called into doubt (the argument goes) when one considers the plethora of past theories that were empirically successful on anyone's standards, but wide of the mark ontologically at the unobservable level. Given the failure of all or most of our past theories, despite their explanatory power, to describe the unobservable, we can reasonably presume that our present theories are likewise inaccurate. Therefore inference to the best explanation can be enlisted towards realist conclusions about the extent of our scientific knowledge, only on pain of an inductively unjustified present-theory chauvinism.

My analysis of the principle of inference to the best explanation in the first chapter suggested that it depends on the establishment of a biconditional relation between truth and explanatory success. I argued that the following argument lay behind the principle:

1. Theory X is the best available explanation.
2. True theories typically will provide the best explanation.
3. Theories that provide the best explanation typically will be true.
4. ∴ Theory X is probably true.
5. ∴ Theory X should be believed to be true.

The argument from the underdetermination of theory by data is best understood as a logical attack on the third premise since it shows, in effect, that there may always be theories which provide equally good explanations, but which cannot both be true. They cannot both be true because they are ontologically incompatible. The historical argument we are now considering aims at undermining both the second and the third premise, and does so through inductive means.

The best recent exposition of the historical objection to scientific realism comes from Larry Laudan. He maintains that a comprehensive sampling from the history of scientific theory shows, on the one hand, that many theories which had little empirical support in their time we now consider to have been substantially correct, and on the other hand, that an even greater number of theories can be found which enjoyed great empirical success, but which we now believe to be false in their postulations. As examples of the former sort of case, Laudan cites Proutian atomic theory, 18th century chemical atomic theory, and the Wegenerian theory of continental drift. Instances of the second sort of theory include the crystalline spheres of ancient and medieval astronomy, the phlogiston theory in chemistry, the electromagnetic ether theory, the caloric theory of heat, and so on. And this list of false theories, Laudan tells us, could be continued *ad nauseum* so as to indicate that there are strong inductive grounds against the realist claim that theories which garner extensive explanatory success typically will be true. These two sorts of historical cases are intended by Laudan to sever both links (premises (2) and (3) above) in the biconditional between explanatory success and truth that supports the realist's adherence to the principle of inference to the best explanation, and her contention that a realist epistemology is warranted by the empirical success of theories. Laudan writes:

> The inescapable conclusion is that insofar as many realists are concerned with explaining how science works and with assessing the adequacy of their epistemology by that standard, they have, thus far,

> failed to explain very much. Their epistemology is confronted by anomalies that seem beyond their resources to grapple with.[10]

I think that the scientific realist need not be concerned about the first sort of historical case, which is supplied as a counter-example to premise (2) above that true theories typically will provide the best explanation for the given domain of phenomena. For even without this premise, the realist conclusion that we should believe those theories that provide the best available explanation will follow if premise (3) is true. That is, even if it is wrong to expect true theories to be empirically successful, the fact that successful theories typically will be true is enough to justify the inference to the best explanation. The realist can admit it as an historical possibility that at any given time the class of empirically successful theories will constitute a sub-class of the class of true theories, and hence that one cannot base the presumption of explanatory success on truth, while still maintaining that empirically successful theories can nonetheless be presumed true. This is all that is required to support the inference from explanatory power to truth, which is, after all, the foundation of the inference to the best explanation. In what follows I will therefore concentrate on the significance of the putative counter-examples to the claim that successful theories typically will be true.

Before considering possible responses available to the realist to the apparently anomalous character of history as far as her central thesis is concerned, I first want to dispense with a certain misguided objection that is sometimes waged against Laudan's historical gambit.

---

10. Laudan (1984a), p. 244

The objection goes as follows. Laudan's argument requires an objective standard of empirical success upon which to base the claim that past false theories enjoyed the same kind of empirical success as the contemporary theories which the realist regards as true or near true. That is, if it is to be held that many now discredited theories of earlier scientific periods exhibited a level of empirical support "no different in kind from that enjoyed by contemporary physical theories"[11], there must be some sense in which empirical success can be compared across theories. And this can only be cashed out in terms of a theory's ability to make correct predictions and (in Laudan's terms) solve empirical problems, i.e., in terms finally of the truth and falsity of a theory's empirical consequences. Therefore, Laudan must allow for a means of determining truth and falsity at the empirical level at least, in order to avoid a relativism of empirical success that would undermine his argument. But, as Newton-Smith puts it,

> on this construal, he cannot maintain the thesis that it is more important to ask whether theories constitute adequate solutions to empirical problems than it is to ask whether they are true, corroborated, or otherwise justifiable within the framework of contemporary epistemology. For in asking whether they provide an adequate solution, we shall have to ask these sorts of questions of the sentences of the theories which are used in the derivation which constitutes the solution to the problem.[12]

11. Laudan (1984b), p. 157
12. Newton-Smith (1981), p. 120

In allowing for an objective notion of empirical success couched in terms of truth and falsity, Laudan is implicitly committed to judgements concerning the truth or falsity of the theories from which the consequences are derived.

However, in admitting an objective notion of empirical success, as Laudan certainly must in order for his argument to work, he is not automatically committed to similar judgements concerning the unobservable aspects of scientific theory. This sort of commitment would follow form a positivistic semantics which reduced the meaning of theoretical statements to that of observation statements. For from that perspective, there would be no difference between theoretical and empirical truth or success. But Laudan's form of anti-realism, like van Fraassen's, is opposed to this sort of semantic reductionism, and can therefore legitimately separate the criteria of the truth of the empirical import of a theory from its truth simpliciter. This has the consequence for Laudan's view that it may well be rational to endorse theories which, for all we know, are radically false. But neither does it rule out the possibility that they are true.[13] The objection misses the point as far as the agnostic brand of scientific anti-realism espoused by Laudan goes, which prescribes the aim of empirical truth and success, separately from judgements about theoretical truth.

What then of the sort of historical cases that Laudan thinks "exhibit the unreliability of inference to the best explanation as a strategy for warranting truth claims on behalf of scientific

---

13. Laudan (1977), p. 126

theories"?[14] I do not want to undertake a detailed historical analysis of the examples brought fourth by Laudan. Let us admit for now that *prima facie* they present serious anomalies for the realist's general inferential move from explanatory success to truth. What avenues are open to the realist for dealing with past theories that were highly empirically successful but false in what they said about the unobservable realm?

On the face of it, there would appear to be two routes that the realist might opt for by way of response. The first would be to assert that these theories in fact lacked the sort of empirical success required to warrant the realist inference to truth. The second response would be to allow that such theories indeed held significant explanatory power, but insist also that they were in some sense approximately true or referential.

The first approach seems hopeless. For one thing, it would be extremely difficult to explain how the empirical success of phlogiston and ether theories differed *in kind* from that enjoyed by contemporary chemical theory. And even if a convincing story could be told in this regard, as some think the methodological anarchists and incommensurablists have done, it would still be incumbent upon the realist to explain why the standards of our present methodology warrant the inference to truth, while those of past theories did not. It is, of course, always possible to bite the incommensurablist bullet and deny both that the transition between theoretical paradigms is rational, and that we can expect the transition to reflect a closer

14. Laudan (1984b), p. 157

approximation to the truth. But this is obviously anathema to the realist. It would seem then that this first response to the historical anomalies must either depend on a suspect chauvinism regarding the kind of success garnered by present theory, or end in a form of methodological anarchism unacceptable to realism in any case.

The second tack open to the realist is to somehow rescue partial reference or approximate truth for the rejected theories, thereby preserving the requisite link between explanatory success and truth. This approach will involve the suggestion that displaced theories in certain scientific fields owe their success to their having achieved a partial or a partially inadequate description of the entities and processes postulated by contemporary theory in the same domain of research. It may be said that successful past theories can be understood as an approximation to present theories by virtue of certain shared functional roles in explanation. The terms of past theories on this account served to isolate some of the causal roles played by the entities of present theories, and for this reason garnered some, though not all, of the explanatory power that contemporary theory possesses.[15] So on this view the theories on Laudan's list did in fact provide a partial or approximate representation of reality. For example, it might be said that the Daltonian atom referred to the atom of present-day physics, that "ether" referred to the electromagnetic field, and that "dephlogisticated air" in the language of Priestley's theory actually

15. Hardin and Rosenberg (1982), p. 614

referred to oxygen.[16] Thus, the ontology of displaced theories retains reference, even though there is nothing in the world that satisfies perfectly their original theoretical characterization.

This way of handling the historical gambit appears more promising than the attempt previously considered. For the idea that it is possible to achieve reference to actual entities through the framework of an unsatisfactory ontology is not entirely counter-intuitive. For example, it seems not unreasonable to suggest that our ancestors' sincere talk of demonic possession involved reference in many instances to what we now recognize as a certain form of schizophrenia. And this interpretation suggests a plausible explanation for the empirical success of displaced theories, relative to that of present day science. That ether theory should have some, but not all, of the explanatory power of electromagnetism is just what should be expected if it is interpreted as an approximation or partial description of the entities and processes identified in electromagnetic theory. In this way realists can explain both why past theories succeed where they do, and failed where they failed. And, in turn, this way of preceding is compatible with a rationalistic reconstruction of the history of scientific methodology, which, as I have argued, cannot be said of the other response to the historical gambit.

But this reply is nevertheless plagued by two major difficulties of its own. First, it relies on the very unclear and controversial notions of partial reference and approximate truth. These concepts currently

16. Kitcher (1978), p. 535

lack detailed formal explication. The realist therefore cannot expect to close the debate by invoking them. While this weakens the force of the reply, it does not wholly dispense with it. For as Newton-Smith points out[17], everyday language is rife with these sorts of notions. We could say, for example, that a certain biographer provided an approximately true description of Joe, though she left out certain of his character traits, and included others that Joe does not really possess. The biographer's description achieves partial reference to Joe even there is no one who answers completely to the description provided in the biography. Still, until the realist can provide the requisite formal analysis, it must be admitted that the burden of proof has not shifted back to the anti-realist.

Secondly, and more importantly, in arguing that the ontology of past theory can be understood as approximatly true relative to contemporary theory, the realist presupposes the truth of contemporary theory. That is to say, the realist's only grounds for claiming approximate truth or partial referentiality for past theories is that their function in explanation resembled in important respects that of theories currently held in the same domain of phenomena. But unless the empirical success of present theory is independently sufficient justification for the belief in their truth, it can not be assumed that the similar, though inferior, explanatory power of past theories stands as an indication of their approximate truth. For all we know, the anti-realist will insist, the functional similarities of ether explanations to electromagnetic explanations makes it that much

17. Newton-Smith (1981) p. 283

more likely that ether theory is false since electromagnetic theory may itself be false. It would seem that unless the realist can show that inference to the best explanation is reliable with respect to current theory, she will not be able to claim that the relatively lesser empirical success of past theories indicates approximate truth. And the pessimistic induction from the history of science strongly suggests that such a present theory application of inference to the best explanation will prove, in the long run, to have been misguided.

There is one counter-argument still open to the realist at this point. Laudan's historical gambit turns on the many cases from the past where empirically successful theories turned out to be false. But the realist may reply that Laudan ignores the other sort of historical case where successful theories turned out to be true, and whose truth can now be determined independently of abductive-style inferences, through direct observation. With regard to these theories, the inference from the best explanation proved to have been a reliable epistemic principle. Examples include Galileo's theory that the moon was mountained, bacterial theory in medical science, and cellular biology. According to Ernan McMullin, the evident reliability of inference to the best explanation in such instances leaves "the burden of proof on the person who denies any connection between truth and explanatory power."[18]

This is too strong. Remember that Laudan's argument is not aimed at revealing a conceptual incoherence or logical flaw in the idea that there is some important connection between truth and explanatory

18. McMullin (1985), p.225

power. Rather, the historical gambit brings inductive evidence against the reliability of epistemological inferences based on the supposed connection. He therefore need not prove that explanatory power and truth are unrelated, only that history does not generally confirm their essential connectedness. And this is sufficient to undermine the realist's application of the inference to the best explanation as a general epistemological principle. In the absence of some explanation on the part of the realist for the fact that the purported relation between empirical success and theoretical truth holds in some historical cases, but not in others, inference to the best explanation can not be assumed to be reliable for any given theory for which there are no independent grounds for thinking it to be true.

Our conclusion from the preceding discussion of the historical objection to inference to the best explanation would appear to be substantially the same as that reached at the end of our analysis of the argument from empirical underdetermination. Understood as an epistemological thesis directed towards the justification of belief in the truth of scientific theories presently held, realism cannot be supported by the principle of inference to the best explanation. In the final section of this chapter, I will argue that a somewhat weakened form of scientific realism can be salvaged through the employment of a second-order inference to the best explanation.

III. The 'No Miracles' Argument

Since Hilary Putnam famously declared in 1975 that the fundamental motivation for realism is that it is the only philosophical position that does not render the success a miracle[19], the 'no miracles' argument has constituted the most popular and persuasive defense of scientific realism. Contemporary realists like McMullin, Leplin, Musgrave, and Boyd have all recently propounded some version of it. I think that the argument is best understood as a more generalized instance of the realist's traditional appeal to inference to the best explanation. However, unlike the theory-specific form of inference to the best explanation that we have so far focused upon, the 'no miracles' argument operates at a meta-level with respect to the epistemological status of scientific theory. Where the theory-specific version took as explanandum the empirical success of some particular theory, the 'no miracles' argument begins with the overall increasing empirical success of science through history. And the inference involved is not to the truth of any single theory presently held, but to the increasing accuracy or verisimilitude of scientific theory as a genuine representation of reality. The basic idea is that a realistic epistemology regarding the progress of science offers the best, and indeed the only, explanation for the undeniable fact that science has steadily increased in explanatory power over the course of intellectual history.

I think that we can break down the 'no miracles' argument in the following way:

1. Science is increasingly empirically successful.

---

19. Putnam (1975), p.73

2. This fact requires some explanation.
3. No anti-realist philosophy of science can provide an explanation.
4. Realism can provide an explanation.
5. ∴ A realist philosophy of science is required.

This is the argument that lies behind the 'no miracles' manifesto of Putnam *et al.* While the argument appears to be valid, its soundness is questioned by the anti-realist. In particular, the third and fourth premises are initially open, respectively, to versions of the underdetermination argument and the historical gambit. In what follows I will contend that, after all, the 'no miracles' argument does not suffer the same unhappy fate as the theory-specific version of the inference to the best explanation when subjected to underdetermination and historical objections of the sort we considered in the last two sections. I will conclude the section with some remarks about the nature of the form of scientific realism that survives.

Construed as an instance of inference to the best explanation, the 'no miracles' argument will need to show that there are no alternative explanations for the success of science that are empirically equivalent to, but incompatible with, realism. That is, it must not be the case that realism is empirically underdetermined relative to some anti-realist account of scientific progress. For if realism, *qua* philosophical explanation, is underdetermined, then premise (3) above is false, and the realist conclusion will be lost.

It is difficult to find an anti-realist attempt to explain the general success of science. Van Fraassen has offered what is probably the best candidate. He casts scientific inquiry in a Darwinian framework, claiming that the increasing success of science should be no surprise, given that empirical adequacy is precisely the criterion upon which theories are, or should be, evaluated. Van Fraassen writes:

> 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 onto actual regularities in nature.[20]

Understood from this evolutionary point of view, the increasing success of scientific theory needs no further explanation than that empirically successful theories typically are chosen over inferior competitors. The realist's appeal to increasing verisimilitude and the like is unnecessary and dismissed as so much metaphysical baggage.

The realist is naturally tempted to object that this does not at all explain the phenomenon in question. For it will be admitted by realist and anti-realist alike that contemporary science is successful because empirically successful theories tend, over time, to be accepted by the scientific community. But this, the realist insists, does not explain why the theories that were accepted by scientists were in the first place more successful than the ones which were rejected. On the Darwinist analogy, we can explain the success of a species by pointing to its capacity for adapting to the environment. But an explanation for the fact that this species was able to so adapt,

20. Van Fraassen (1980), p. 40

and hence survive, where other now-extinct species were not, will require appeal to certain deep structural features it possessed which the others lacked. In the same sense we can explain the success of contemporary scientific theory by citing its adaptive success in its environment--by pointing to the fact that it possesses sufficient empirical strength to warrant its acceptance by scientists. On the other hand, if we want to know why the theories currently held possess the empirical strength they do and have hence survived, while others lacked comparable empirical success and were subsequently rejected, we will need a structural sort of explanation. And the explanation will in this case be realistic: the theories in question owe their greater empirical success to the fact that they are a more accurate reflection of the unobservable world.

Still, it is not clear that a structural explanation is required by way of response to the challenge presented by the 'no miracles' argument. What apparently is in need of explanation is simply that science is increasingly empirically successful. And this can be explained adequately by pointing out that empirical success is precisely the basis upon which scientific theories are selected. Therefore increasing empirical success is no miracle. Indeed, it is exactly what we should expect given continuing adherence to a rational scientific methodology.

Nevertheless, I think there is an improtant aspect of the success of science that cannot be captured by the anti-realist's Darwinist explanation. This is roughly the fact that scienctific theories are successful in ways that do not form part of the basis for their original acceptance. That is, scientific theories normally sustain their empirical success over long periods of application, over sometimes novel phenomena, and in domains of research beyond that of their original application. Science is successful not only in the sense that successful theories are chosen by scientists, but often also in quite unexpected ways--in ways that do not contribute to the justification of their being accepted in the first place. And while it is true that this kind of success in turn constitutes a part of the rationale for their continued acceptance, it remains that at any stage of theoretical development, the success of the theories that are accepted often turns out to be much greater, both in kind and in range, than that which was the primary warrant for their acceptance. In this sense, the often unexpected success of scientific methodology is not explicable merely on the grounds that successful theories survive. For this sort of success plays no part in the justification of their survival; it is always an additional, and unexpected, empirical achievment. I suggest, therefore, that science is successful, and increasingly so, in a respect that cannot be adequately captured by the claim that only successful theories survive. We should conclude that the Darwinian account of scientific progress is at least not as good an explanation as that offered by the realist. Survival can explain only the success of scientific theories that justifies their acceptance, and not the further novel success commonly garnered by those which do survive.

That survival does not explain success can be illustrated further by reference to the actual character of theory-shift and competition. In general, new theories are accepted for their ability to overcome empirical anomalies that plagued their predecessors. In this sense they are born successful, while their survivability is a function of their future success (or progressive problemshift as Imre Lakatos puts it[21]) relative to competing theories that would hope to supplant them. And more importantly, this survivability is at best only an indirect function of success since degenerative research programs for a multiplicity of reasons might well be maintained for long periods of time, even in the face of extremely limited empirical success. The claim that science is increasingly successful because only the more successful theories survive in the competitive jungle is thus disconfirmed by the frequent survival of unsuccessful theories.

This brief glance at scientific practice suggests a more general problem with the anti-realist's Darwinian explanation of success. Might it not be argued that science is not properly termed a biological phenomenon at all in the sense that, say, procreation is a biological phenomenon? Rather science is an essentially human artefact, that requires intellectual capacities and values that only we possess. Unlike the mechanisms of natural selection, scientific methodology is a thoroughly rational activity. In particular, theory choice is determined entirely by the *will* of scientists, and is in this sense a purely teleological phenomena. But as Darwin repeatedly urged, natural selection must be understood in non-teleological terms. This fact alone should make us wary of any simple analogies between scientific progress and the evolution of natural species, and

21. Lakatos (1970)

hence of any attempts to explain the success of science in terms a 'scientific survival of the fittest'.

I am not sure whether or not there are any other kinds of anti-realist explanations for the increasing empirical success of science. In any event we will see in the next chapter that a more promising strategy on the part of anti-realism is to deny that the realist's explanation serves to refute an agnostic attitude towards scientific theory. But I will conclude for now that there are no completely satisfactory anti-realist accounts of progress that would show the underdetermination of the realist explanation, that science is providing an increasingly accurate picture of the world at both the observable and unobservable level.

Besides the underdetermination thesis, the historical induction of Laudan constituted the other important objection against the inference to the truth of contemporary scientific theory from its explanatory success. Does the inferential strategy characteristic of the 'no miracles' argument fall to the same kind of objection? It would appear not. For the historical gambit undermined the theory-specific form of inference to the best explanation by means of counter-examples. But the 'no miracles' inference is to the overall increasing representational accuracy of theory, from the increasing empirical success of science. It is not clear what would count as an historical counter-example to the reliability of this sort of inference, if for no other reason than that the explanandum includes all of scientific history. The sort of case required to count as an anomaly for the inferential strategy behind the 'no miracles' argument would have to be one which suggested what is evidently not true: that since the time of Aristotle, science has not generally progressed in terms of explanatory strength.

The reason the 'no miracles' argument is not subject to Laudan's historical gambit is that the argument is not aimed at justifying belief in the truth of any particular theory--past, present or future. It therefore cannot be waged as an objection against the second-order inference to the best explanation that it involves an epistemic leap of faith. In adopting the 'no miracles' approach, the realist takes a lesson from history and dissociates himself from the claim that some specific theory can be warrantedly judged true on the basis of its empirical success. In so doing, he is able to avoid the charge of an unjustified confidence in the truth of present theory that emerges from the historical induction, and premise (4) of the above formulation of the 'no miracles' argument remains intact.

Nonetheless, in moving to a meta-version of the inference to the best explanation, the realist buys refuge from the underdetermination and historical difficulties at a significant cost. The kind of scientific realism that follows from the 'no miracles' argument confines its epistemological claims to the scientific enterprise generally; it lacks the resources to justify belief in the truth of any particular theory. This attenuated form of realism will therefore contend that there is good reason to think that science is giving us through its theories an increasingly accurate picture of the nature of the reality that lies beneath the observable, but will share with anti-realism an agnostic stance towards the theories actually currently held by scientists.

Still, to concede this much is not to collapse realism into the position of its anti-realist opponent. An important difference remains between the two approaches as regards the aim of science. The anti-realist typically will insist that the only goal that science should be expected to achieve is that of providing a true account of the phenomena. Even the attenuated form of realism that emerges from the 'no miracles' argument goes further than this, claiming that the empirical success attained by science provides sufficient ground for the *further* belief that it is genuinely approaching a more accurate and complete description of the world in terms of both its macro and its micro-structure. Therefore, understood as a thesis about the extent of the increase in our knowledge that science has provided, realism is still incompatible with anti-realism. And if the 'no miracles' argument is sound, realism gives us what anti-realism apparently cannot: a philosophical explanation for the astounding success that science has had, and continues to have.

In the next chapter, I will introduce a final ingenious objection to the inference to the best explanation which, if successful, refutes even this weakened form of realism arrived at via the 'no miracles' argument.

CHAPTER THREE:
ANTI-REALISM AND EXPLANATION

Introduction

In the first chapter I argued that the realist - anti-realist debate in contemporary philosophy of science centers on epistemological questions, rather than metaphysical or semantic ones. In particular, the concern is with the legitimacy of the realist's employment of the principle of inference to the best explanation towards the justification of belief in the truth of the unobservable claims of scientific theory. In the second chapter I discussed the two most important arguments that have been proposed by anti-realists against inference to the best explanation: the argument from empirical underdetermination and the historical gambit. I concluded that together these arguments serve to undermine the realist's inference from the empirical success of any particular theory to its truth, or approximate truth. But in the final section of the second chapter I argued further that a non-theory-specific second-order version of inference to the best explanation, in the form of the 'no miracles' argument, did not succumb to these anti-realist objections. The 'no miracles' argument moves from the overall increasing empirical success of science to the conclusion that science is in fact achieving an increasingly accurate description of the unobservable realm, thus preserving a characteristically realist thesis about the aim of science.

In this chapter I introduce a far more radical anti-realist objection, which suggests that the 'no miracles' argument must either beg the question against anti-realism, or remain insufficient to establish any uniquely realist conclusions. Unlike the underdetermination and historical arguments that were discussed in chapter two, the objection that I will present is not aimed at revealing any particular flaw in the inference to the best explanation. Rather the anti-realist in this case objects that the realist inferential principle presupposes--and must presuppose if it is to have any force--a view of explanation that the anti-realist does not accept.

I. Inference and Belief

The usual objection raised by the anti-realist upon the introduction of the 'no miracles' argument is that, in retreating to a meta-version of the inference to the best explanation, the realist cannot hope to do any better than she did with her original theory-specific application of the principle. For the basic pattern of inference, an abductive one, is the same in both cases. With the theory-specific version, the inferential path is from the explanatory success of a particular scientific theory to its truth, while the 'no miracles' argument attempts to move from the explanatory success of a particular philosophical theory--realism--to its truth. The premises are different, but the argument form is the same. But if, as I have argued, inference to the best explanation fails at the level of ordinary scientific practice, the realist surely cannot hope to succeed by the employment of the same principle at the meta-level. I have claimed that the central issue distinguishing realist from anti-realist attitudes towards scientific knowledge is the question of the epistemological legitimacy of the abductive inference to the best explanation. But the 'no miracles' argument is just another manifestation of the adherence to this sort of inferential strategy, and therefore cannot be invoked towards a resolution of the debate, especially given the failure of this strategy at the ground level of actual scientific reasoning. As Arthur Fine expresses it, "his [the realist's] success at the methodological [meta] level can be no better than his success at the ground level. If he fails there, he fails across the board."[1] [my inserts]

But understood in this way, the objection has no weight. While it is certainly true that the legitimacy of inference to the best explanation is central to the controversy, it is no objection to an argument for realism that it takes an abductive form, unless one's intention is to rule the possibility of a scientific realist epistemology out of court. To avoid begging the question herself, the anti-realist will need either some specific argument against the second-order, 'no miracles'-type inference to the best explanation, or alternatively some general reason for rejecting *any* form of abductive inference to realist conclusions. This is especially apparent given what was said in section III of the last chapter. For there I tried to show that the particular difficulties raised by the anti-realist to undermine the theory-specific version of inference to the best explanation (underdetermination, historical gambit) did not apply to, and hence were ineffective against, the 'no miracles' argument. Therefore, it will not suffice simply to dismiss the 'no miracles' argument as yet another instance of the realist's inferential strategy (it certainly is

1. Fine (1984), p. 88

that), since it was the tenability of certain premises of the original formulation of the argument, not the form of argument itself, that came under anti-realist fire. The objection to the 'no miracles' argument we are now considering therefore seems to turn on a conflation of the notion of the soundness of an argument, with that of the validity of a general argument form. So, in the absence of some general problem with inference to the best explanation, the realist need not be troubled by the observation that the 'no miracles' argument is another application of her central epistemological principle.

Yet there is a way of construing this sort of objection according to which the success of the 'no miracles' argument can be seen to rest on a fundamental circularity in the realist's inferential strategy. The principle of inference to the best explanation is founded on the idea that something's providing the best available explanation for a given state of affairs provides sufficient warrant for the belief that that something is true, or exists. Thus, at the theory-specific level the realist urges that the relative explanatory success of, say, electromagnetic theory, justifies that belief that the theory is true. A similar justificatory relation is supposed to hold at the meta-level between the overall increasing empirical success of science and the belief that science is achieving an increasingly accurate description of the unobservable realm: the realist view of the aim of science. The truth of electromagnetic theory is the best explanation of its extraordinary empirical success; the truth of realism is the best explanation for scientific progress in general.

In chapter two, I noted the particular difficulties that plagued the first sort of inference, and argued that they did not affect the second. But the anti-realist is not opposed to inference to the best explanation as a means of justifying belief only in certain isolated cases; rather she is against it in principle. The anti-realist simply does not accept the basic realist tenet that explanatory power indicates truth, and she therefore finds fault with inference to the best explanation across the board. Therefore, if the anti-realist would separate explanatory power and truth completely, the realist must not offer as grounds for belief in realism its role in successful explanatory stories, on pain of begging the central epistemological question: are we justified in believing the best explanatory hypotheses to be true?

At this point, a dilemma emerges for the realist. Recall our analysis of the 'no miracles' argument in the last chapter. I argued that it was sound--that the premises were immune to anti-realist counterexamples. The conclusion was that to the extent that the empirical progress of science required an explanation, only realism, the view that science is providing an increasingly better description of the world through its theories, can provide it. But for the anti-realist, it still does not follow that realism is true. That is, it does not follow from the fact that realism provides the best available explanation for the progress of science that we are justified in believing that scientific theories actually constitute increasingly accurate descriptions of the unobservable. Consider in this regard the following passage from <u>The Scientific Image</u>:

> I say that Newton could explain the tides, that he had an explanation for the tides, that he did explain

> the tides. In the same breath, I can add that after all this theory is incorrect.[2]

The same may presumably be said of the realist explanation of scientific progress. Therefore, even if the 'no miracles' argument is sound, the conclusion it arrives at is not sufficient to defeat the alternative anti-realist conception of the aim of science.

The realist will, of course, deny that Newton's theory *really* explains the tides. For, as we will later see, realists construe explanation as a simple objective relation between theory and phenomena, and hence believe that only 'true' theories can provide 'true' explanations of the phenomena. With this van Fraassen would agree, except that he sees explanation as an irreducibly pragmatic notion, to which the terms 'true' and 'false' do not properly apply to begin with. From an anti-realist perspective there is therefore no inconsistency in saying that a false theory such as Newton's *really* explains. This aspect of the debate will be examined in more detail in the next section. For now we can at least see that the 'no miracles' argument will be ineffective against a anti-realist who denies that explanations are truth-related in any simple way.

The argument could, of course, be rendered adequate by the addition of a further premise which said that we are justified in believing the best explanation for scientific progress to be true. But this would beg the question against the anti-realist conception of scientific explanation, which, again, is that explanation is a theoretical virtue independent of judgements of truth or falsity. The realist is on

---

2. van Fraassen (1980), p. 99

the horns of the following dilemma: either the 'no miracles' argument is too weak to establish any uniquely realist conclusions, or it begs the question by assuming a theory of explanation that the anti-realist patently rejects.

In the second section of this chapter I explore in more detail an anti-realist view of explanation. But first I want to mention a standard objection to the wholesale rejection of inference to the best explanation which allows the anti-realist to escape the apparently realist conclusions of the 'no miracles' argument. If the anti-realist wishes to dismiss all instances of abductive inference to the best explanation, she is in danger of rendering illegitimate a good deal of the inferential reasoning that is fundamental to scientific methodology. For it is a commonplace that theories are sometimes formulated and selected primarily on the basis of their ability to explain a given range of phenomena. Consider the following idealized case of theoretical inference:

1. Phenomenon O
2. If theory T then O
3. $\therefore$T

This is a simplified example of abduction, and appears to be the way in which a good deal of scientific theorizing generally proceeds. So if the anti-realist is to renounce abductive inference totally, he must dismiss it as practiced in actual theory formulation. This done, it is not clear that there would remain any legitimate scientific methodology at all. Richard Boyd accentuates the disastrous consequences of the rejection of abductive reasoning with the following remark:

> The rejection of abduction or inference to the best explanation would place quite remarkable strictures on intellectual inquiry. In particular, it is by no means clear that the students of the sciences, whether philosophers or historians, would have any methodology left to study if abduction were abandoned.[3]

Given what was said in the second part of chapter one about the nature of abduction and inference to the best explanation, I think that there is an easy response to this objection available to the anti-realist. Recall my distinction between abduction and inference to the best explanation. I identified the former as a legitimate methodological principle, and the latter as carrying an extra epistemological premise aimed at justifying belief in the truth of the theory that survives the abductive procedure. It seems clear that the anti-realist is opposed only to the epistemological element. He accepts the abductive method as a valid means of serving the proper aim of science: the attainment of empirically adequate theories. What he rejects is the further claim that their methodological success provides grounds for the belief that they are true. For according to the anti-realist view of explanation, the explanatory power of a theory is a pragmatic warrant for its acceptance, but never independent evidence for its truth. He can therefore preserve abduction as a legitimate tool of scientific methodology, while still insisting that inference to the best explanation and the 'no miracles' argument presuppose an epistemological view of explanation which he does not accept.

---

3. Boyd (1984), p. 67

However, the problem re-emerges in a more serious way concerning the positive claims of some anti-realist philosophies of science. For example, van Fraassen construes the aim of science merely in terms of the *acceptance* of empirically adequate theories, as opposed to the *belief* in true theories, where a theory is empirically adequate if and only if gives a true account of the phenomena. But even acceptance involves belief--the belief that the theory is empirically adequate. And given that a theory is empirically adequate only if it saves *all* the phenomena, past, present and future, the belief involved in theory acceptance necessarily goes well beyond the available evidence.[4] So even the anti-realist conception of the aim of science must include, as a condition for theory acceptance, a belief which is not strictly determined by observation. Where the realist infers from increasing empirical success to the belief in increasing verisimilitude, the anti-realist infers from empirical success thus far to the belief in empirical adequacy. If like the former, the latter is justified by an inference to the best explanation, then the original anti-realist objection to the 'no miracles' argument--that explanatory power is not grounds for belief--collapses. *Tu cocques.*

I should say first of all that it is by no means essential to anti-realism that it explicate the aim of science in terms of empirical adequacy. Larry Laudan, for instance, identifies the aim of science as the acceptance of theories which have shown themselves to be more effective problem-solvers than available alternatives. But in this case

4. van Fraassen (1980), p.12

acceptance involves only the belief that they have thus far solved a relatively high number of problems; not that they have solved, or will solve, all problems.[5] Thus, on this anti-realist view of the aim of science, acceptance need not include any beliefs which go beyond the available evidence one currently holds for the problem-solving adequacy of a theory.

But even van Fraassen's attitude regarding theory acceptance can be reconciled with the rejection of inference to the best explanation. The important question is this: is the belief in empirical adequacy properly understood as being justified by an inference to the best explanation? It seems clear, first, that a judgement of empirical adequacy has the same character as an empirical generalization. That the theory in question will save *all* the phenomena is inferred from the fact that it has *so far* been completely empirically correct. Therefore the belief is founded on an inductive inference. This is not the case with the 'no miracles' argument, which does not merely generalize from a finite number of observations, but introduces a new sort of term (increasing verisimilitude) as an explanation above and beyond all inductive conclusions about the empirical success of scientific theories. The belief in empirical adequacy, by contrast, is derived from a generalization across known phenomena. And while this inductive inference may be questionable in its own right, it is certainly different from the realist's inference from the explanatory power of realism to the belief in its truth.

This notwithstanding, Gilbert Harman argues that inductive

5. See Laudan (1977), p. 111

arguments are actually best understood as constituting a sub-class of the class of arguments based on inference to the best explanation.

> If we think of the (inductive) inference as an instance of inference to the best explanation, we can explain when a person is and is not warranted in making an inference from 'All observed A's are B's' to 'All A's are B's'. The answer is that one is warranted in making this inference whenever the hypothesis that 'All A's are B's' is a better, simpler, more plausible, explanatory hypothesis than is the hypothesis, say, that someone is biasing the sample to make us think that 'All A's are B's'.[6]

Applying this analysis to the belief involved in theory acceptance, we would say that the hypothesis that a theory is empirically adequate is the best explanation for the observed fact that it has so far saved all of the phenomena, making it an instance of inference to the best explanation after all.

There are two related points that may be made in response to this. First, I am not sure that it is correct to characterize the justification of inductive inference in this way. Imagine drawing 99 black marbles in turn from an opaque jar which is known to contain a total of 100 marbles. There appear to be two possible explanations for our having drawn 99 black marbles: that they are all black, and that all but the last one chosen are black. The former appears to be a better explanation and so, according to Harman, we are warranted in our inductive conclusion that they are all black. But why exactly is this a better explanation? Obviously it is because of the extremely

6. Harman (1965), p. 91

low likelihood that the single non-black ball would be chosen last. But this fact is *already* sufficient to justify the inductive conclusion that all of the balls are black. That is, an inductive generalization counts as the best explanation only by virtue of already qualifying as a legitimate inductive inference. So it cannot be the case that induction is warranted by inference to the best explanation since the same probablistic facts that identify the generalization as the best explanation justify the inductive inference in the first place.

My argument here has a characteristically anti-realist flavour. That 'All A's are B's' is the best explanation for the observed phenomena does not provide any grounds for believing it beyond whatever grounds we already have. For whatever reasons we have for thinking that a certain empirical generalization is the best explanation are themselves sufficient for thinking that the generalization holds. We should therefore conclude that it is incorrect to classify inductive generalizations as instances of inference to the best explanation, since it is not the explanatory power of the generalization that provides the initial justification for the inductive inference. Whether or not the belief in empirical adequacy can ever be justified in its own inductive terms is an interesting, but separate, question. It would seem at least that the belief is compatible with the rejection of inference to the best explanation as used by the scientific realist.

This brings me to the second point that I want to make about the difference between the belief involved in theory acceptance and the belief in increasing verisimilitude: the latter *must* rely on an explanatory inference because it is about unobservables. The belief in empirical adequacy, being directed entirely at (oberved and unobserved) observables is logically equivalent to belief in an indefinitely long conjunction of statements about phenomena. Consider van Fraassen's example of the belief 'there is a mouse in the wainscoting'. Given that the mouse is itself an observable thing, "'there is a mouse in the wainscoting' and 'all observable phenomena are *as if* there was a mouse in the wainscoting' are totally equivalent; each implies the other (given what we know about mice)."[7] Similarly, the belief that a theory is empirically adequate is logically equivalent to the belief that all of the observable phenomena are *as if* the theory was empirically adequate.

Compare this to the realist's belief in increasing verisimilitude. This belief is not equivalent to a conjunction of statements about the phenomena. That is, there is a logical gap between the belief that all of the phenomena are *as if* science was providing increasingly accurate descriptions of reality, and the belief that science is actually doing so. The belief in increasing verisimilitude of the 'no miracles' argument therefore necessarily goes beyond the power of induction, and must rely for its justifcation on its purported explanatory power. On the other hand if what I argued earlier is correct, the anti-realist's belief in empirical adequacy does not depend on explanatory inference for its justification. This further suggests that the inference to increasing verisimilitude is of a different, and much more suspect, epistemological species than the inference to empirical adequacy.

This shows that an anti-realist of van Fraassen's ilk is not subject

7. van Fraassen (1983), p. 21

to the objection that his view of the aim of science must, no less than that of the realist, include a belief which goes beyond all the evidence. For being wholly about observables, the belief in empirical adequacy stands or falls according to the canons of inductive logic (whatever they are). And hence it does not rely for its justification on considerations of explanatory power. Furthermore, this all appears to be consistent with a thorough-going anti-realist epistemology, particularly as regards the abandonment of the sort of explanatory inferences upon which the 'no miracles' argument must depend.

Our discussion so far in this chapter strongly suggests that the debate between realists and anti-realists concerning the aim of science reduces finally to opposing views about the epistemological status and significance of explanation. In the next section I will investigate and contrast the theories of explanation possessed by each side of the controversy. I will attempt to isolate the central disagreement, without arguing for one side or the other. But I do hope in this way to show how even this fundamental opposition leaves open the possibility for a version of scientific realism, which I shall call 'entity realism', that is compatible with both the realist's, and the anti-realist's, views about the nature and role of scientific explanation.

II. Truth and Explanation

Beginning with the pioneering work of Hempel and Oppenheim in the late 1940's[8], questions concerning the nature and role of scientific explanation have attained an increasingly important position within discussions in the philosophy of science. There is now a huge wealth of literature covering every aspect of the subject, and containing every conceivable philosophical opinion. I do not intend in what follows to canvass the various analyses of scientific explanation that have been recommended, nor will I even defend or criticize any particular view of the matter. Rather, I will restrict my discussion to a description of the fundamental characteristics of, and basic motivations for, an anti-realist conception of the epistemological status of explanation. My aim is to show that the anti-realist's contention that explanatory power is a purely pragmatic virtue of scientific theory is the philosophical basis for his rejection of inference to the best explanation. We need to understand the anti-realist's view of explanation in order to evaluate the objection introduced in the last section to the 'no miracles' argument: that any such 'explanationist' defence of realism must either beg the question, or remain insufficient to decide between the opposing attitudes towards the aim of science. This section will also serve as background for the next chapter, where I will argue that the pragmatic aspects of scientific explanation do warrant agnosticism where causal explanations are concerned. This fact, I will argue, leads immediately to entity realism.

Some philosophers of an anti-realist bent have, in their disdain for the unobservable, gone so far as to deny that science is in the

8. Hempel and Oppenheim (1948)

business of providing explanations at all. This was certainly the opinion of Duhem, for example. Duhem understood explanation as the attempt "strip reality of the appearances covering it like a veil, in order to see the bare reality itself."[9] This realistic view of explanation led him to reserve its pursuit for metaphysical speculation, thus restricting the aim of science to the mere classification of phenomena. Some philosophers have interpreted van Fraassen's more startling declarations on this subject--"There are no explanations in science"[10]-- as indicating a similar view. We shall see however that such bald statements must be understood in relation to his general theory of explanation. For what van Fraassen means to deny is that science gives us explanations in the way that realists (and Duhem) understand the notion.

But in any case, whatever explanation is, anti-realism does not wish to deny it an important role in scientific inquiry. Van Fraassen writes, "A theory is said to have explanatory power if it allows us to explain; and this is a virtue."[11] The anti-realist arguments of both Larry Laudan and Nancy Cartwright are based on the assumption that explanatory power is at least one aim of scientific theorizing. But they try to show that this theoretical virtue cannot be a reliable indicator of truth. Laudan's historical induction is meant to undermine the assumed connection between explanatory success and truth by drawing from the long list of past theories which, though

---

9. Duhem (1954), p. 7
10. van Fraassen (1977), p. 150
11. van Fraassen (1980), p. 97

successful, we now believe to be false.[12] Cartwright, on the other hand, maintains that theories of high explanatory power can only be had at the expense of facticity.[13] In either case, the motivation for their anti-realism depends on the idea that explanation is among the aims of science.

What anti-realism *does* oppose is the belief that explanatory power can be an epistemological virtue of theory--that the search for explanation is a genuine independent means of increasing our knowledge of the world. In particular, the anti-realist denies that the explanatory power of a theory gives us reason to think it true, over and above any reason we have for believing it to be acceptable, empirically adequate, instrumentally reliable, and the like. This claim that explanatory power does not provide independent grounds for belief has two possible motivations. First, it may be argued that explanatory power actually increases the likelihood that a theory is false. This is the opinion of Cartwright, who claims that there is a trade-off at the level of theoretical laws between explanatory efficacy and truth. Secondly, and alternatively, it is argued that explanation has nothing whatsoever to do with epistemological issues, that explanatory power is a virtue of scientific theories entirely distinct from judgements of truth and falsity. And therefore extensive explanatory power constitutes a reason for accepting a theory, but never a reason for believing it. This, roughly, is the position of van Fraassen.

---

12. Laudan (1984)
13. Cartwright (1983)

Since I discuss in some detail the views of Cartwright and their relation to entity realism in the next chapter, and since van Fraassen has offered what is probably the most comprehensive anti-realist theory of explanation, in what follows I concentrate almost exclusively on his work.

Most recent forms of scientific realism have relied, implicitly or explicitly, on some version of the principle of inference to the best explanation. But in order for it to favour realism, the principle has in turn been supported by a characteristically realist interpretation of explanation. The basic realist intuition in this regard is that there is some connection between explanation and truth. One way of cashing out the connection is by making truth part of the definitional criteria of explanation: nothing can count as an explanation unless it is true. On this view, to the extent to which we consider a theory to be explanatory, we must *ipso facto* presume it to be true. This leads immediately to realism. For it implies that to the extent that science aims at finding explanations, it aims at discovering theories that are true. So realism must be the correct view of science on this account.

But this is obviously an untenable position. For one thing, it would follow from this logical conception of the relation between truth and explanation that once we have decided that a particular theory will explain the phenomena, its truth would be guaranteed by simple entailment. But this would imply either that no now-rejected theories of the past explained anything, or that they were all true. While the latter option is clearly contradictory, the former represents an abuse of both scientific and ordinary language. For we certainly *do* want to say that Newton had an explanation for the tides, that caloric theorists had an explanation for heat transfer, that Copernicus' theory of concentric orbits around the sun explained the apparent motions of the heavenly bodies, and so on, even though none of these theories is in fact true. Similarly, we may say that senility is a possible explanation for some of President Reagan's statements, even though we do not think that he is senile. It appears that what we should say of such explanations is that they are false or unsatisfactory, not that they are not explanations. But on the view we are presently considering, which says that explanations require true premises, the term 'false explanation' is contradictory. And 'unsatisfactory explanation' makes no sense, unless taken to mean 'false explanation', in which case it is contradictory.

Indeed the entire realist strategy of invoking inference to the best explanation depends on the idea that the connection between truth and explanation is of an epistemological, not a logical, variety. For the notion that the superior explanatory power of a theory stands as *evidence* for its truth presupposes that we recognize explanation prior to, and independently of, truth. This would not be the case if truth was made part of the definition of explanation. So the realist must admit that explanation *simpliciter* is distinct from truth, if she is to be able make sense of concepts such as 'true', 'false', and 'the best' explanation, which are central to her defence of realism. What the realist does want to claim is that explanatory power stands as epistemic justification for the belief in truth, that better explanations are more likely to be true explanations, even given the logical possibilty that a theory of extensive explanatory power turn out to be false. Explanation suggests truth, but does not imply it.

The realist's contention that explanatory power is grounds for belief derives its force from a particular model of scientific explanation. This is the 'standard' model, according to which explanatory power is a simple function of the relation between theory and phenomena. Roughly, a theory is judged to have greater or less explanatory power in proportion to the number and extent of true empirical consequences it will generate, considered in conjunction with certain accepted auxiliary statements. We can see how this will tend to favour realism. For on this model, a theory T counts as the best explanation by virtue of the fact that the phenomena which actually obtain are more to be expected given the truth of T, than the truth of some other theory. Therefore, it is more reasonable to believe in all cases the best explanation than the alternatives which have more false, or less true, consequences. This is the basis for the realist's claim that explanation is an epistemic virtue.

We have already seen that this way of reasoning cannot defeat agnosticism where is is the truth of any particular theory that is being inferred (see first two sections of the previous chapter). But we have also seen that increasing verisimilitude can be identified as the best explanation for the increasing empirical success of science by the employment of a second-order abduction. And if explanation is a simple two-term relation between theory and fact, it is a better explanation by dint of being more likely to be true given the fact of empirical progress, than the alternative supposition, that science is not so progressing. So long as explanation is construed simply in terms of the empirical consequences of a given theory, where an explanation is better if it has more true and fewer false consequences, realists can always trade on the intuitively plausible idea that true theories are more likely to be successful in this respect than false ones. Hence the justification of inference to the best explanation. It would seem therefore that if the anti-realist hopes to deny the epistemological import of explanatory power, and thereby block inference to the best explanation, she will need a radically different conception of the structure of scientific explanation.

This, in fact, is just what van Fraassen offers. He maintains that explanation is a pragmatic virtue, and hence not a ground for belief. He identifies explanation as a three-term relation between the theory that is supposed to explain, the phenomena to be explained, and the context of the demand for explanation (where the context is determined by the interests of the person who happens to be asking for an explanation). Explanatory power is not a relationship *sui generis* between theory and fact, to be evaluated simply on the basis of the empirical consequences that a theory generates. Rather, explanatory power is relative to the interests and values of individuals. It may therefore naturally give reason to prefer a certain theory, but never reason to believe it. If van Fraassen is correct in thinking that the value of an explanation is necessarily a function of pragmatic concerns, then the explanatory power of a theory does not make it more likely to be true than an alternative, only more practical to accept. Hence the illegitimacy of inference to the best explanation.

Van Fraassen argues for his pragmatic theory of explanation against the standard model, primarily by trying to show that the classic anomalies that plague the standard model result directly from the failure to recognize the essential context-dependency of explanation. In The Scientific Image, he evaluates several attempts at providing a philosophical account of explanation in terms of an isolated relation between theory and phenomena. For brevity's sake, I will discuss only his criticisms of what is probably the twentieth century's most influential view of the structure of explanation: Hempel's covering law model. My selectivity should be harmless since van Fraassen maintains that the problems he locates apply to all non-pragmatic models of explanation.[14] I will put-off discussion of van Fraassen's pragmatic interpretation of causal explanations until the next chapter.

Under Hempel's covering law model, explanations take the form of arguments. Where non-statistical theories are concerned, e.g. relativity theory, the argument is deductive. A theory is said to explain a given phenomenon just in case the phenomenon can be derived from the theory in conjunction with certain relevant initial conditions. This is the D-N (deductive-nomological) version of the model. As for irreducibly statistical theories, e.g. quantum mechanics, the argument-form is inductive. A theory is said to explain just in case the theory, together with certain relevant initial conditions, confers a high probability on the phenomenon. This is known as the I-S (inductive-statistical) version of the covering law model.

There are two classic difficulties which, according to van Fraassen,

---

14. van Fraassen (1980), p. 111

show that the covering law model presents neither sufficient nor necessary conditions for explanation. The first problem is that the covering law model seems to make explanations symmetrical. The standard example concerns the explanation for the length of the shadow cast by a flagpole. According to the D-N model, the length of the shadow is explained by the height of the flagpole since the former can be derived from the latter, together with the assumption of certain initial conditions, including the angle of elevation of the sun, and a basic trigonomic formula. However, we can also deduce the height of the flagpole from the length of the shadow, assuming the same initial conditions. But we presumably do not want to say that the height of the flagpole is explained by the length of the shadow.

The same difficulty confronts the I-S version of the model. Imagine that there is a very high correlation between the possessing of a certain chromosome pair (X-Y), and having blue eyes. In this case, a person's possessing that chromosome pair will confer a high probability on his having blue eyes, and hence will explain his eye colour according to the I-S model. But it will also be true that there will be a high probability that a person will possess the X-Y pair, given that they have blue eyes. But again, we do not want to say that the eye colour explains the chromosome pairing. The covering law model does not capture the basic asymmetry of explanation, and as a result, allows too much to count as explanation. So fitting the covering law model is not sufficient to qualify something as an explanation.

The second problem suggests that the covering law model is not even a necessary condition, since it excludes apparently legitimate instances of explanation. The most famous illustration of this involves a person who asks her doctor why she has the disease paresis. The doctor, knowing that no one contracts the disease unless they have a history of latent untreated syphilis, replies, "because you have latent syphilis, which was allowed to go untreated". This appears to be a satisfactory explanation. But of all the people that have latent untreated syphilis, only 1% ever contract paresis. Therefore, the person's having latent untreated syphilis neither implies, nor makes highly probable, her eventually contracting paresis, and hence does not qualify as an explanation according to the covering law model. So the model is not a necessary condition for explanation: it allows to little to count as an explanation.

This is, of course, all familiar territory. The problems with the covering law and related models have been well documented. My concern, however, is not so much with the problems themselves as with van Fraassen's efforts to resolve them through his pragmatic theory of explanation. As I have said, where philosophers traditionally have understood explanatory power as a simple diadic relation between theory and phenomena, van Fraassen includes a third factor, context. And van Fraassen claims that it is the failure to recognize the role of context in the evaluation of explanatory power that has led to the major difficulties.

The context of explanation is roughly the particular subjective interests of the person who is demanding explanation. The context of explanation may vary relative to what is apparently the same request for explanation. Consider the following request for explanation: "Why did Adam eat the apple?". There are at least three legitimate, but distinct, sorts of answers available, depending on the context in which the question is asked. For example, a theologian may wish to know why Adam ate the apple, as opposed to not eating it. In this case the appropriate explanation would be something like, "because, being of flesh, he was unable to overcome temptation". Or a biblical scholar may, by the same question, be asking for an explanation of why it was an apple (that is, the fruit of the tree of the knowledge of good and evil) that Adam ate, as opposed to some other sort of food. An explanation here would perhaps recount something of the religious significance of tree-grown fruit in the time that the book of Genesis was composed. Finally, a sociologist interested in gender roles in religious myth may be demanding an explanation for why it was Adam that was tempted by Eve, and not vice-versa. In this case a satisfactory explanation may adduce information concerning attitudes towards women in early Hebrew culture.

The point is that in each instance, what counts as an adequate explanation is a function of a certain contrast-class. That is, before we are in a position to evaluate T as an explanation for P, we must have in mind a certain range of possible alternatives to P. A demand for explanation, according to van Fraassen, always has the underlying form of a request for an explanation of a particular event relative to a contrast-class, or range of alternatives. But as we saw in the Adam example above, the content of the contrast-class is determined by the context--the particular interests of the person demanding explanation. Thus the theologian wished to know why Adam ate the apple, as opposed to not; the biblical scholar was interested in his eating the apple as opposed to some other food; and the sociologist wanted an explanation for the fact that it was Adam, and not Eve, who ate the apple. (*I apologize to those of the Christian religious tradition for treading rather heavily on this biblical passage in Genesis*).

The context-dependency of the contrast-class points to a solution to the kind of problem illustrated by the paresis story. For there are two basic kinds of explanations that may be appropriate as explanations relative to the person's question, "why do I have paresis?" depending on the context. If she is interested in why she, of all people, has contracted paresis, the appeal to her history of untreated syphilis is a legitimate explanation. If, on the other hand, she is interested in the fact that only she among her fellow syphilitics contracted paresis, an entirely different sort of explanation is required. Thus, van Fraassen contends that we can only judge the goodness of an explanation relative to a triadic relation between explanans (phenomenon), explanandum (theory), and a contextually-determined contrast-class. And this third factor is entirely contextual in the sense that, for a given explanation, neither the phenomenon to be explained, nor the theory which is to do the explaining, alone provide the relevant information concerning the peculiar interests of the person who is demanding the explanation. We can therefore evaluate explanations only by tacitly including an element that is irreducibly a pragmatic function of the interests and desires of the person who is requesting explanation. Hence, the extent to which we find an explanation powerful is always partly a reflection of our interests, independently of the truth or falsity of the theory to which we appeal as explanation.

The other major problem--explanatory asymmetry--can be cleared up, van Fraassen claims, by focusing on a second important contextual factor in the evaluation of explanations: relevance. Relevance is that property of an explanation which identifies a particular event as the most pertinent of the various factors that play a contributing role in bringing about the phenomenon to be explained. The context-dependancy of relevance can be illustrated by the following example. A father asks his son, "Why is the porch light on?" What counts as a satisfactory explanation in such circumstances will depend on the context in which the question is posed. If the father is interested in the particular human expectations or desires which led up to the depressing of the switch, an answer along the lines of, "Because we are expecting company", would be suitable. On the other hand, were the son, citing different factors that led up to the state of affairs of the light being on, to reply "Because the porch switch is closed, and electricity is reaching the bulb through that switch", we would probably accuse him of impudence. This is because in the given context in which the father's question is posed, the factors cited as explanation are irrelevant. But as van Fraassen points out, we can imagine a context in which the situation is exactly reversed. For example, the father is re-wiring the house. Upon noticing the porch light on, and fearing he has caused a short-circuit bypassing the porch light switch, he asks the same question. This time, the son's appeal to the technicalities of electrical circuitry would be relevant. In this context, it is this sort of explanatory factor that interests the father. Therefore, the relative value of explanations is dependent upon the relevance of certain possible explanatory factors, which is in turn a function of the peculiar interests and desires of the person who is demanding explanation.

If van Fraassen is correct in this, we have the basis for a solution to the problem of explanatory asymmetry. For if explanation is context-relative in this sense--if context is an essential element of explanation--then asymmetries should be explicable on account of a radical shift in context. In other words, it should be possible to construct situations wherein the explanatory relation between two events can be reversed as a consequence of a change in context. To this end, playing on the famous flagpole-shadow case of asymmetry, van Fraassen devises a set of circumstances in which the length of the shadow is explanatorily relevant to the height of the flagpole, given the peculiar interests of a person who asks, "why is this flagpole so high?"[15] I will not go into the details of the piece of philosophy of science fiction van Fraassen uses to illustrate the contextual dependency of explanatory relevance. The main point for van Fraassen's account is that the asymmetry of explanation is no problem so long as we appreciate the significant influence of context on our evaluation of explanations. For the asymmetry of explanation can always be preserved relative to a given context. But if we exclude the contextual factors, and construe explanation as a simple diadic relation between explanans and explanandum as in the covering law model, it is difficult to see how the asymmetry can be

---

15. van Fraassen (1980), pp. 132-134

adequately captured.

As I said at the beginning of this section, I do not intend to place van Fraassen's pragmatic theory of explanation under heavy analytic scrutiny. Although I will criticize some aspects of his account in the following chapter, my primary objective in what precedes has been to illustrate the philosophical basis for an anti-realist view of explanation. We have seen that van Fraassen's strategy is to emphasize the importance of the pragmatic context in which demands for explanation are couched. In particular he has tried to show that the extent to which we find an explanation compelling or adequate is in part determined by a certain contrast-class and relevance relation in mind. But these contextual factors are determined neither by theory (explanans) nor by phenomena (explanandum); rather they are a function of the particular interests and desires of the person who is asking for explanation. Hence, there can be no judgements about the explanatory power of a theory, as such. This, I submit, is the idea behind van Fraassen's cryptic and frequently quoted remark that "there are no explanations in science".[16]

We can now see how this all supports the anti-realist's contention that explanatory power does not itself provide ground for belief, and her subsequent objection to the 'no miracles' argument--with its employment of inference to the best explanation--that it must beg the question against anti-realism. For the realist is committed to the idea that the belief in realism is justified over anti-realism by virtue

16. van Fraassen (1977), p. 150

of the fact that the former provides a better explanation for the success of science. But if van Fraassen is correct, this purported explanatory power of realism is a reflection of its ability to satisfy certain of our desires and interests in a certain context, but does not give us reason to believe it to be true over and above the reasons we have for believing it to be empirically adequate. Therefore, the anti-realist may conceivably admit that realism is an empirically adequate theory of science, that it also has the pragmatic virtue of extensive explanatory power, but remain agnostic as concerns the actual aim of science--and all this without inconsistency. At the meta-level realism stands to scientific progress as atomic theory stands to the unobservably small. They are both empirically successful, possibly empirically adequate, and perhaps should be accepted as such. But the fact that they carry with them great explanatory power is at best further grounds for accepting them, independently of questions of their truth or falsity.

CHAPTER FOUR:
ENTITY REALISM

Introduction

It is a main tenet of this thesis that the standard arguments for scientific realism are explanationist in form. The most significant attempts at providing a rational defense of realist doctrines concerning the epistemological status of scientific theories, or the aim of scientific inquiry, have relied--tacitly or otherwise--on some version of the principle of inference to the best explanation.

We have already seen that this explanationist strategy faces major barriers where it is the truth or approximate truth of current scientific theory that is offered as the best explanation for its extensive empirical success. In particular, I have argued that the problem of empirical underdetermination and the historical gambit present fatal difficulties for any theory-specific application of the principle of inference to the best explanation.

These sorts of problems do not, however, affect a second-order application of the principle, where a realist thesis about the aim of science is inferred from its being the best explanation for the overall increasing empirical success of science. But we have found that this meta-version of inference to the best explanation, characteristic of the 'no miracles' argument, must either presuppose a view of explanation that the anti-realist rejects, thus begging the question, or remain insufficient to establish any uniquely realistic conclusions. For the anti-realist understands explanation as an irreducibly pragmatic feature of theories; explanatory power furnishes grounds for acceptance, but never for belief. So the apparent explanatory power of the philosophical theory of realism as regards the empirical success of science can, as far as anti-realism is concerned, at best only suggest that realism is empirically adequate. But it does not provide independent evidence that realism is true.

Van Fraassen's anti-realism about explanation is more radical than the agnosticism he prescribes for scientific theories themselves. For he maintains that the claims of theories are literally true or false, but that we cannot be justified in believing them to be true, or to be false. On the other hand, he seems to think that explanations do not have truth-values at all: they are not true-or-false, only pragamtically better-or-worse. So in the fashion of a deeper Dummetian kind of anti-realism, the principle of bivalence is rejected for explanations. This leads to the conclusion that inasmuch as theory is considered to possess great explanatory power, this in itself is no evidence for its truth. Explanation simply has nothing to do with these sorts of epistemological judgements.

In the final chapter I aim to show that this radical form of anti-realism about explanation cannot be sustained where the explanations in question are causal. Following Nancy Cartwright, I argue that when one accepts a causal explanation, one is thereby committed to the existence of the cause. It does not make sense to assert that X explains Y, by virtue of X being the cause of the effect Y, but still remain agnostic regarding the existence of X. So while it may be admitted that explanations are generally neither true nor false, it cannot be denied that the entities which function in causal explanations either do or do not exist. It is this way of reasoning that motivates the most recent brand of scientific realism to emerge: theoretical entity realism.

In what follows, I examine the two most important arguments, from Nancy Cartwright and Ian Hacking, that have been offered in favour of entity realism. I argue that, ultimately, both arguments must rely on a version of the realist standby--inference to the best explanation. I will end the chapter with some conclusions about the nature of entity realism, and its relation to the broader problem of providing a philosophical account of scientific progress.

I. Cartwright

The most prominent theoretical entity realist summarizes her view of explanationist arguments for realism with the following slogan: "no inference to the best explanation, only inference to the most likely cause."[1] Cartwright argues that in order for the fundamental laws of physics to function adequately in good explanatory accounts, they *must* be false. There is, she maintains, a necessary trade-off at the level of theoretical laws between explanatory power and facticity. And given that explanation is a primary aim of scientific inquiry, Cartwright is led to adopt an anti-realist view of scientific theories. But a good deal of scientific reasoning, especially in experimental contexts, is from effect to cause, where the effect is a particular phenomenon and the cause is some unobservable theoretical entity. And Cartwright urges that this form of reasoning is legitimate. So she is a realist concerning the entities science postulates to account for

1. Cartwright (1983), p. 86

certain phenomena.

Before getting into the details of Cartwright's argument, it will be helpful to draw a natural distinction between two general types of scientific explanation. Following the terminology of Ernan McMullin,[2] I will call these *nomothetic* and *structural* explanations. We offer a nomothetic explanation of a particular event or phenomenon when we cite it as falling under a regularity or generalization, usually known as a law. Thus, if asked "Why is that crow black?", we may explain "All crows are black". Or we may explain the exponential rate of electromagnetic decay of an atom by appealing to Schroedinger's mathematical equations. Structural explanations on the other hand are those whereby we explain the peculiar behaviour or qualities of an object or system by referring to aspects of its deep structure. For example, we may explain the colour of an object by pointing to its molecular structure which reflects incident light frequencies at a certain spectral angle. Or we may explain the explosion of a sealed jar that is heated by referring to the expansion of the gases it contains. Structural explanations typically appeal to the constituent elements of the object or phenomenon whose behaviour or observable qualities are to be explained.

With both nomothetic and structural explanations, the explanans may be either observable or unobservable. Newton's inverse square law of gravitation, for example, was thought to explain the motion of bodies by virtue of being a true abstract empirical description of observable phenomena, without appeal to hidden forces. On the other

2. McMullin (1978)

hand, the complex equations of contemporary quantum theory, which govern and explain the behaviour of quantum systems, are directed entirely at the unobservable. Likewise with structural explanations. We may, for example, explain a person's high blood pressure by citing the cholesterol build-up on the walls of his main arteries. Conversely, the violent result of mixing two chemicals is explained by the unstable interaction of their constituent molecules.

I mention this because I want to focus primarily on explanations that make recourse to the unobservable. Since we are interested in the question of the extent to which the role of certain laws or entities in successful explanations argues for their truth or existence, repectively, I will restrict my analysis to explanations with explanans we have no independent emprirical grounds to believe in. Therefore, with nomothetic explanations I will concentrate on the so-called fundamental laws of physics, from which the empirical laws, or 'phenomenological' laws, as Cartwright calls them, are supposedly derived. Similarly, structural explanations that will concern me are those in which the entity or process invoked as explanans is unobservable.

The most important difference between nomothetic and structural explanations is that the latter are usually causal. When we explain the track in the cloud chamber as the path of an electron, or the sun's emission of radiant energy as the result of hydrogen fusion at its core, we are appealing to causes. On the other hand, when we explain the rate of decay of a radioactive substance in terms of abstract quantum mechanical equations, or the force between two charged bodies in terms of Coulomb's Law, we are relying not on causes, but rather on universal generalizations which particular instances of the relevant phenomena obey. For Cartwright this distinction between structural and nomothetic explanations--between explanation by cause and explanation by nomological instantiation--is crucial. For, while she agrees with van Fraassen that truth is external to explanation, she believes that causal explanations have a built-in existential component. Hence, nomothetic explanations may be completely adequate, yet false. (She actually thinks their falsity to be a natural consequence of their explanatory power in many cases.) But in accepting a causal-structural account, we are *ipso facto* committed to the existence of the cause. Therefore, Cartwright's realism includes only the entities and processes that function in causal explanations, not the fundamental laws that furnish successful nomothetic explanations.

Before examining the realist half of Cartwright's position, I want to discuss briefly the motivations for her anti-realism about theoretical laws. She distinguishes first between two basic kinds of laws which operate in theoretical science: phenomenological and fundamental laws. Roughly, phenomenological laws are descriptive generalizations that successfully summarize and organize the known phenomena in a particular domain. They are directed at actual, as opposed to abstract or ideal, physical processes and behaviour. As Cartwright says, "Phenomenological laws describe what happens".[3] Examples include Kepler's laws of planetary motion and the specific laws used in aerodynamics and electrical circuitry engineering.

3. Cartwright (1983), p. 2

Fundamental laws, conversely, are highly abstract and general, rather than concrete and descriptive. They cover and unify a large range of phenomenological laws under a single set of explanatory principles, usually mathematical formulae. Examples of fundamental laws are Maxwell's and Schroedinger's equations, and the basic formulae of special relativity. It is important to notice that the distinction between phenomenological and fundamental laws is not made along the observable-unobservable line, but is based rather on theoretical function. Phenomenological laws are concrete, specialized, and descriptive of actual processes; fundamental laws are abstract, unifying, and highly explanatory.

Cartwright has no principled objections to phenomenological laws. She thinks they are, or at least may be, true generalized descriptions of how the world behaves. On the other hand, she maintains that the fundamental laws of physics patently do not state the facts, and thinks that their falsity is essential to their explanatory function. In How the Laws of Physics Lie, three main arguments are offered against the 'facticity' view of fundamental laws, all of which focus on the tendency of important explanatory strategies to detract from veracity.

First, she notes that most fundamental laws carry *ceteris paribus* clauses. That is, they are understood to govern the behaviour of actual phenomena only under ideal conditions, where all other possibly relevant factors not mentioned by the law are held constant. Read literally, without the *ceteris paribus* qualifier, the laws are almost all false, for the simple reason that "there are no exceptionless quantitative laws in physics".[4] In fact, the best candidates are known to have anomalies--as Popper observed, all theories are born refuted.

If, on the other hand, we consider the fundamental laws with the *ceteris paribus* clauses intact, they are true only on certain extremely rare occasions--when the conditions are ideal. The fundamental laws are therefore either simply and obviously false, or only rarely true. But Cartwright argues that *ceteris paribus* generalizations are nevertheless essential to the explanatory role of fundamental laws, for almost all explanations we consider to be strong, both within and outside of scientific contexts, carry *ceteris paribus* assumptions. Hence, the appendage that 'everything else is equal', which we know is almost always false, is nevertheless indispensable to fundamental laws fulfilling any significant explanatory function in science.

The second argument centers on the common use of approximation procedures in the calculation of predictions about the behaviour of actual phenomena from the basic equations of fundamental laws. Theoretical realists tend to understand this detour from strict rigorous deduction through approximation as a practically necessary, and very marginal, departure from the exact truth. If we were to undertake the long and complex calculations from the fundamental laws, the truth of the resultant statements about the actual phenomena, the realist contends, would be preserved as a matter of entailment. In opposition to this view, Cartwright argues that approximation actually takes us closer to the truth about the phenomena than would be the case if the rigourous calculations were

---

4. Cartwright (1983), p. 46

all worked out. Therefore, the fundamental equations cannot themselves be literally true of the world. For example, the exponential decay law, which enjoys a wealth of empirical support, can only be derived from the basic equations of quantum theory with the aid of significant approximations. Similarly, predictions about the actual state of an electrical circuit of any complexity will be radically off the mark when strictly derived from the relevant fundamental laws. Accurate predictions are obtainable only through systematic approximations.[5] In either case, the form of approximation is dictated by the recognition of certain causal factors within the system that are not identified by the fundamental laws. Strictly speaking, the basic equations will not themselves generate any true statements about the actual physical systems.

It may be objected that the fundamental laws are meant to characterize only idealized situations, and that it is always possible in principle to work into them all of the other relevant factors that guided the approximations, thereby obtaining equations which *will* make accurate predictions about the phenomena. But as Cartwright points out, this strategy will result in "a longer and longer list of complicated laws of different forms, and not the handful of simple equations which would be fundamental in a physical theory."[6] By 'internalizing' in this way the approximations required to generate correct predictions, the basic equations become increasingly

---

5. For an extended discussion of these examples, see Cartwright (1983), pp. 107-118
6. Cartwright (1983), p. 112

particular and directly descriptive. And with this loss of simplicity and generality goes the over-arching and broad explanatory power that characterizes them as fundamental laws of nature. As with *ceteris paribus* generalizations, the role of approximation in actual scientific practice indicates a trade-off of explanatory power and facticity at the level of fundamental theoretical laws.

Cartwright's final argument that the laws of physics lie concerns explanation by combined forces. She notes that there is no single fundamental law which will govern the behaviour of any actual physical object or system exclusively. Instead, actual physical bodies are typically subject, at any given time, to a number of laws from various theoretical domains. For example, the behaviour of most bodies will be governed at any given time by both the law of universal gravitation and Coulomb's law (the law which gives the force between the electrical charge of bodies). The actual behaviour of these bodies must therefore be explained by the combination of two forces, and will not be truly described by either of the laws singly. "No charged objects will behave just as the law of universal gravitation says; and any massive objects will constitute a counter-example to Coulomb's law."[7] Because we do explain the behaviour of objects by combining laws in this way, we necessarily allow that the laws we use patently do not state the facts about any actual activity in the world.[8]

---

7. Cartwright (1983), p. 57

8. There is an interesting and complex issue about whether or not it is plausible to regard the forces identified by combined laws as actually describing the behaviour of the system in question. An

It may be objected that while neither of the laws used to explain the behaviour of physical systems subject to combined forces manage to truly describe any actual behaviour, it is possible to construct a 'super-law', a kind of mathematical amalgamation of the two original laws, which *will* be true of the phenomena that actually obtain. Cartwright's reply to this is two-fold. First, even when such covering-laws are available, they may not provide very satisfying explanations. Consider the demand for explanation, "Why is the force between the two bodies p?" A super-law combining gravity and the force of electrical charge can apparently offer only: "Because the bodies are subject to covering-law Z." This is because the super-law, being merely a quantitative conjunction of the original laws, fails to unify the laws relative to a broader background theory in a way that would illuminate the behaviour of the object in question. For this reason, explanation by so-called super-laws often seem like Moliere's *vis dormativa* explanations. Though perhaps true, they are not very explanatory.

On the other hand, in some cases it is possible to construct super-laws which do not just piece together the particular influences involved, but provide a more abstract unifying description that fits in with a wider body of theory. In this case, Cartwright argues, the super-law may well be informative, but will miss an important feature of explanation by composite laws. Specifically, they miss the

---

affirmative answer has the consequence of making the resultant force of their combination a "mathematical fiction". See Creary (1981), for an extended discussion.

fact that the reason the predictions made by the super-law obtain in the actual world is because the forces identified by the original laws and unified by the super-law are operative in the given situation. In other words, super-laws can never wholly *replace* explanation by component forces, so long as we want to say that the behaviour of certain bodies is a consequence of the combined operation of distinct fundamental laws. "To understand how the consequences of the unified laws are brought about would require separate operation of the law of gravity, Coulomb's law, and so forth; and the failure of facticity for these contributing laws would still have to be faced."[9]

Cartwright's anti-realism shares with van Fraassen's a view of explanation that departs radically from the standard covering-law model. According to the covering-law model, a theory is said to explain a certain phenomenon when the phenomenon can be shown to follow from the fundamental laws of the theory. Cartwright's "simulacrum" account of explanation places far greater emphasis on the place of theoretical models as intermediate between the simple mathematical formulae of the theory, and the complex and diverse phenomena to be explained. According to Cartwright, the fundamental laws of the theory explain the phenomena by attempting to fit them into models whose form is determined by the basic equations of the theory. The models help us to understand the phenomena, says Cartwright, by allowing it to be 'seen' through the broader, more elegant, mathematical framework of the theory. So, "to

9. Cartwright (1983), p. 71

explain a phenomenon is to find a model that fits into the basic framework of the theory and that thus allows us to derive analogues for the messy and complicated phenomenological laws which are true of it."[10]

But Cartwright argues that the models used in this explanatory function stand only as 'simulacra' for the actual phenomena: a good idealized representation, but not a literally true reflection of the actual physical processes. The fundamental laws of the theory are true of the objects in the model, but not of the actual phenomena. For the properties of the model will often include obvious idealizations--infinite potential, friction-less planes and the like--which perhaps approach, but do not genuinely describe, reality. In other cases, certain properties of the model will be pure fictions, employed for reasons of practical convenience rather than exact representation. This all serves the pursuit of explanation, but it is explanation in a characteristically anti-realist sense. The theoretical models explain the phenomena by organizing and logically classifying them. They explain by setting the apparently disparate phenomena in the broad and elegant framework of a unified theory, rather than by revealing the actual phenomena as special cases of the true laws of nature. And this end, Cartwright maintains, is admirably served by high-level physical theories, even at the expense of the facticity of their fundamental laws and equations.

Nonetheless, Cartwright's anti-realist conception of explanation, unlike that of Van Fraassen, is not intended to apply universally.

---

10. Cartwright (1983), P. 152

There is a form of explanation commonly pursued in science which she thinks does provide grounds for belief: causal explanation. This brings us to the realist half of Cartwright's philosophy of science, her defence of theoretical entity realism. There are two basic components to this defence. First, she argues that unlike nomothetic explanation or explanation by fundamental laws, causal explanations require existential commitment. She here adopts an explanationist strategy, based on inference to the best *causal* explanation. Secondly, she argues that unlike theoretical explanations, which rely entirely for their justification on inference to the best explanation, causal explanations "have an independent test of their truth: we can perform controlled experiments to find out if our causal stories are right or wrong."[11] Thus, experiment furnishes independent evidence for the view that inference to the best causal explanation is a legitimate form of explanatory inference.

In my discussion of Ian Hacking, whose arguments for entity realism center almost entirely on experimental intervention, I will argue that such experimental arguments depend for their validity on the legitimacy of inference to the best causal explanation. I contend that the experimental arguments pull in the direction of entity realism only on the assumption that the causal version of the explanationist strategy is sound to begin with. My examination of Cartwright's motivation for espousing entity realism will therefore concentrate only on inference to the best causal explanation.

Cartwright's explanationist defence of theoretical entity realism is

11. Cartwright (1983), p. 82

straightforward and intuitively compelling. Whatever one's theory of explanation, it is inconsistent when reasoning from effect to cause to accept a certain explanation, yet deny that the causal factor invoked to explain the effect exists. This is not to say that the use of causal explanations requires us to hold with an attitude of *certainty* that the cause exists. For it may turn out that we were mistaken, that the phenomenon was actually caused by something else. But in this case, the entity realist claims, it must be admitted that the object we originally identified as the cause does not really explain the phenomenon. Unlike theoretical explanations which, if van Fraassen is right, have an irreducibly pragmatic dimension, causal explanations are right or wrong in a quite objective sense. There is only one correct causal explanation of a particular phenomenon: the actual cause of the effect to be explained. Non-existing causes simply do not explain anything. Therefore, inasmuch as we are confident that we have found a causal explanation for a certain phenomenon, we are thereby committed to the idea that we have identified an actually existing object or process. Moreover, inasmuch as much as we are justified in believing we have found the best causal explanation, we are thereby justified in believing that the entities functioning in the explanation exist: hence the legitimacy of inference to the best causal explanation.

Consider a couple of illustrations from Cartwright, beginning with an instance of nomothetic explanation. Imagine that most of the camelias you have planted in the rich hot soil of your back yard die within a week. Your neighbour's gardener informs you that camelia roots will not take in soil which is above a certain temperature. This looks like the right explanation, for you have taken care to make sure they were properly planted in most other respects. Still, the law-like generalization you appeal to as explanation--that camelias will not survive in hot soil--is not strictly true. In fact, a few of your camelias did survive. And while it is likely that there is an unrecognized factor distinguishing the surviving camelias, it remains that, barring omniscience, we accept the 'hot soil' explanation as legitimate, even though we *know* that the law-like explanans is not true. Nomothetic explanations do not require true explanans to warrant their acceptance as genuine explanations.[12]

Consider now a single isolated causal explanation. A lemon tree which you have planted in an oak barrel earlier in the year begins to show signs of disease: the leaves are yellowing and dropping off. Your neighbour's gardener suggests that stagnant water may have collected at the bottom of the barrel, starving the tree of vital nutrients. Upon investigation, you find that indeed the oak barrel is full of foul water. The gardener's explanation is correct. On the other hand, had you not found the stagnant water, you would no longer accept the explanation as adequate, for then the supposed cause of the lemon tree's disease would be known not to exist. Unlike the previous nomothetic explanation, this causal explanation requires for its legitimacy that the explanans be true, i.e., that cause of the tree's dying actually is the stagnant water. We cannot reasonably accept a causal explanation without an accompanying existential commitment

12. Example from Cartwright (1983), p. 51

towards the cause invoked.[13]

In scientific reasoning, the pursuit of causal explanation coincides with a realist conception of the aim of science. Whatever methodological procedures (abduction and retroduction) guide our identification of the best explanation in instances of reasoning from effect to cause, it equally provides evidence for the belief in actually existing causes. For one cannot consistently affirm that science is partly in the business of giving causal explanations, and that nevertheless we have no reason to think that science is likewise aiming at knowledge of the unobservable entities that typically stand as explanans in those causal explanations.

It may be objected at this point that all this talk about causal explanations runs roughshod over a long philosophical debate about the metaphysical stautus of causation. At least since the time of Hume, many philsophers--especially those of an empiricist bent--have argued that causal relations in nature must finally be construed in terms of a contingent association of events. While I think entity realists will naturally tend towards a metaphysically realist position about causation, it may be possible to side-step (or at least bracket) the entire question. For the important point for entity realism is not that instances of reasoning from effect to cause parallel an objective relation in the world, but rather that such reasoning can be epistemologically distinguished from explanation by appeal to laws. The entity realist may remain metaphysically neutral in the dispute about the reality of causal powers, while insisting that a good deal of

13. Example from Cartwright (1983), p. 91

reasoning in science does appear to be from cause to effect, and that these causal explanations are not always reducible in such contexts to explanation by nomological association. So while it remains an open question as to whether or not causal relations can be distinguished from nomological relations in purely metaphysical terms, the entity realist insists that they can nonetheless legitimately be distinguished epistemologically, in terms of the sort of inferences each allows us to draw from their use in explanations. To this extent, the entity realist falls back on the argument above that isolated causal explanations have an inherent existential import which is lacking in nomothetic explanations.

We have now identified the basic argument for the realist half of Cartwright's theory of science. She believes in the legitimacy of inference to the best causal explanation, and hence believes in the theoretical entities that function in causal explanations. Recall from the previous chapter that the most fundamental anti-realist objection to explanatory inference is that since explanation is an essentially pragmatic, context-relative feature of scientific theories, explanatory power is no indication of truth. Van Fraassen's pragmatic theory of explanation is supposed to apply to all forms of explanation, both within and outside of science. It is intended to undermine inference to the best explanation, not just in certain forms, but across the board. What then of Cartwright's adherence to inference to the best causal explanation? Are causal explanations subject to van Fraassen's pragmatic interpretation?

Cartwright depends in order for her entity realist argument to go through on the idea that causal explanations are objective in the sense that for a given effect or phenomenon, it will in principle be possible to isolate the 'real' cause independently of the context and peculiar interests of the person who happens to be seeking explanation. This is necessary in order to avoid an underdetermination problem whereby, given a certain perspective, various incompatible factors can all be seen as *the* cause of the phenomenon in question. For as we saw in chapter two, the tenability of the principle of inference to the best explanation is contingent upon there not being a number of empirically equivalent yet ontologically distinct accounts of the phenomena. The same goes for Cartwright's causal inferences. Causal explanations must be objective if they are to sustain epistemological inference.

Against the objectivity of causal explanation, van Fraassen makes much use of an example that is now standard in the literature. In the example, we are looking for a causal explanation for a person's death in a single-vehicle accident. Van Fraassen quotes N.R. Hanson who writes,

> There are as many causes of x as there are explanations of x. Consider how the cause of death might have been set by a physician as 'multiple haemorrhage'; by the barrister as 'negligence on the part of the driver'; by a carriage-builder as 'a defect in the brakelock construction'; by a civic planner as 'the presence of tall shrubbery at that turning'.[14]

According to van Fraassen, this sort of example indicates that what is identified as the correct causal explanation will depend on the

14. Hanson (1958), p. 54

interests of the person demanding explanation. He writes,

> the salient features picked out as 'the cause' in that complex process, is salient to a given person because of his orientation, his interests, and various other peculiarities in the way he approaches or comes to know the problem - contextual factors.[15]

It is suggested that this will also be so with causal explanations in scientific contexts. Therefore, the scientist's identification of a certain factor as 'the cause' will always be partially a function of that scientist's pragmatic context. He will have good reason to accept the explanation he arrives at, but not to believe it.

But the example exploits a conflation of several distinct explananda. It would appear that each of the people involved is explaining something quite different. This can be illuminated by appealing to a point which van Fraassen himself makes, that all requests for explanation presuppose a certain range of alternatives--a contrast-class. Recall the request for explanation discussed in chapter 3, "Why did Adam eat the apple?". With this sort of question, what counts as a possible explanation will depend on a certain contrast-class of the form "Why did Adam eat the apple as opposed to . . . ?". We can apply this insight to the carriage accident example, so that each person can be understood to be providing answers for different requests for explanations. Thus, the doctor explains why the person died, as opposed to survived. The barrister explains why this person died, as opposed to those with similar carriage models on

15. van Fraassen (1980), p. 125

the same turn. The carriage-builder explains why this particular carriage crashed, while others did not. And finally the town-planner explains why the accident occurred at this particular turn, rather than at equally sharp corners in the city. The example trades on a vast ambiguity in the suggested request for explanation, "Why did the person die?" Once we have clarified the various distinct sorts of explanations that may be demanded by such a question, depending on the particular contrast-class, it is no surprise that the question should allow of four different causal explanations. For the question will itself allow of four different interpretations suggesting four different phenomena to be explained.

It is significant that van Fraassen fastens on an example from everyday life in order to make his case for the context-dependency of causal explanations. For a convincing argument can be made showing that in the realm of precise scientific reasoning it is clear to all what would count as a legitimate causal explanation for the phenomenon at hand. In scientific contexts, explanations are pursued within the confines of a particular theoretical framework, where there are strict parameters binding the range of possible explanatory hypotheses. For example, consider the request for a causal explanation for the track in a cloud chamber. The contrast-class is quite clear: there not being a track. All other factors are held constant. I suppose a taxpayer may offer as an explanation, "Because the government has provided enough funds to this university for it to buy the apparatus to produce such odd effects". But the scientists would recognize such an explanation as irrelevant. For the question is clearly not why there is a track in the cloud chamber as opposed to there not being a cloud chamber at all.

The point is that in general scientific inquiry is precise enough to eliminate from causal explanations the ambiguity that plagues questions like "Why did the person die?" in the carriage example. Therefore, I think it is reasonable to suppose that scientists looking for a causal explanation for a given phenomenon will share an understanding of the contrast-class. As van Fraassen puts it, "the way he approaches or comes to know the problem" will generally be identical for scientists working in clearly defined experimental situations. There is little reason to think that, in this sense, the explanations they provide will be severely infected by context and interest. And we can have some confidence that it will be clear to all what form 'the cause' could take.

Van Fraassen considers this sort of objection, and responds in the following way.

> It might be thought that when we request a *scientific* explanation, the relevance of certain hypotheses, and also the contrast-class, are automatically determined. But this is not so, for both the physician and the carriage-builder are asked for a scientific explanation. The physician explains the fatality *qua* death of a human organism, and the carriage-builder explains the fatality *qua* carriage crash fatality. To ask that their explanations be scientific is only to demand that they rely on scientific theories and experimentation, not wives tales.[16]

16. van Fraassen (1980), p. 129

I think that this ignores the precise and extremely controlled conditions of scientific inquiry, and experimentation in particular. It is important in scientific contexts to attempt to make it as clear as possible what it is that is supposed to be accounted for in order to rule out irrelevant or trivial explanations. This, I think, is exactly Cartwright's point when she writes,

> We make our best causal inferences in very special situations--situations where our general view of the world makes us insist that a known phenomenon has a cause; where the cause we cite is the kind of thing that could bring about the effect and there is an appropriate process connecting the cause and effect; and where the likelihood of other causes is ruled out. This is why controlled experiments are so important in finding out about entities and processes we cannot observe. Seldom outside of the controlled conditions of an experiment are we in a situation where a cause can legitimately be inferred.[17]

The sorts of contextual factors van Fraassen cites--contrast-class and relevance--are clarified and held fixed as much as possible in experimental situations. So there is little ground for the claim that the causal explanations accepted by scientists should so radically context-dependent as to render epistemologically suspect the inference from effect to cause.

II. Hacking

I now want to consider the main argument of Ian Hacking, the

17. Cartwright (1983), p. 6

other major proponent of entity realism. Though he allies himself closely in some respects with Cartwright, his general strategy is very different. While Cartwright's defense of entity realism is primarily inferential, Hacking focuses almost exclusively on our ability to directly manipulate unobservable entities in experiment. He claims that our best grounds for belief in theoretical entities is to be derived, not from their powerful explanatory function, but rather from our actual use of them in experimental contexts. In particular, he argues that the best evidence we can have for the existence of, say, electrons, is the fact that they can be used as a tool in experiment to test hypotheses from distinct theoretical domains. He thus marks a shift in emphasis, as regards scientific realism, from the *representation* of theoretical entities in explanatorily successful theories and models, to our *intervention* in actual physical processes. Referring to an experiment in which niobium balls are sprayed with electrons to alter their charge and produce 'free' quarks, Hacking summarizes his interventionist justification for entity realism with the following slogan: "If you can spray them, then they are real."[18]

Hacking is a severe critic of traditional explanationist arguments for scientific realism. He disposes of inference to the best explanation both at the ground level of simple inference to the truth of particular empirically successful theories, and at the meta-level, where a realist conception of the aim of science is offered as the best explanation for the overall increasing empirical success of science. His objections to explanationism are not particularly insightful or novel. He suggests

---

18. Hacking (1983), p. 23

that the best we can hope for in terms of an explanation for the success of science is a quasi-evolutionary account centering on our species' essential rationality. "If you must have an explanation for the success of science, then say what Aristotle did, that we are rational animals that live in a rational universe."[19]

Hacking's most basic objection to explanationist strategies seems to be that explanation is not an epistemic notion. Rather he classifies it as a kind of feeling of intellectual satisfaction which is relative to individual interests and the historical and psychological circumstances of the moment. Echoing closely van Fraassen's anti-realist view of explanation, he claims that "Explanations are relative to human interests. . .there are times when we feel a great gain in understanding by the organization of new explanatory hypotheses. But the feeling is not a ground for belief."[20]

This noted, I want to argue that despite his avowed anti-explanationism, Hacking's interventionist arguments for entity realism ultimately depend for their strength on the legitimacy of inference to the best causal explanation. While his arguments do provide a strong intuitive pull in the direction of realism, they can at best compliment, not replace, the more standard explanationism of Cartwright.

As I have said, the basic motivation behind Hacking's entity realism is a recognition of the use that theoretical entities are put to in experimental design. We can no longer reasonably doubt the

19. Hacking (1983), p 57
20. Ibid, p. 53

existence of unobservable theoretical entities which come to play an integral role both in the manipulation of actual physical processes, and in the construction of new experimental devices which use well-understood properties of the entities to interfere in other "more hypothetical parts of nature."[21]

I think it will be admitted that practising experimental scientists are no doubt indeed convinced of the reality of theoretical entities used in these sorts of ways. Whether or not they are actually justified in this conviction is another matter. For the committed anti-realist philosopher will insist that Hacking's argument, so baldly stated, simply begs the question. Like Descartes' "cogito ergo sum", Hacking's interventionist slogan, "If you can spray them, then they are real", assumes from the beginning exactly what is at issue: the reality of unobservable theoretical entities. While it is certainly true that objects we can manipulate must be real, the central question remains as to whether or not we can be justified in the belief that we are actually manipulating unobservable entities in such circumstances.

Hacking is no doubt aware of this logical point. But it does not concern him, for he is skeptical about the very possibility of an epistemological justification of scientific realism. His aim is not to convert an intransigent anti-realist through any putative canons of inferential reasoning. Rather, he wants the anti-realist to feel the natural realist pull that the experimentalist feels as a result of his intervention in actual physical processes. Hacking urges that "the

21. Hacking (1983), p. 265

whole family of issues about realism and anti-realism is mickey-mouse, founded on a prototype that has dogged our civilization, a picture of knowledge *representing* reality."[22] The claim that his interventionism begs the question against an anti-realist like van Fraassen, whose position is grounded in a very traditional brand of skeptical epistemology, will not be very distressing for Hacking. For he believes that the traditional debate will only be resolved when we cease to think of unobservable entities solely as theoretically-ensconced *representations* of reality, and begin to focus more attention on their important function in actual experimental contexts. This sentiment is the basis of remarks like, "The final arbiter in philosophy is not how we think, but what we do".[23]

The anti-philosophical bravado notwithstanding, I think it is clear even in Hacking's own arguments that a rational justification of entity realism, whether or not the doctrine is supported by the intuitive pull of interventionism, must ultimately rely on inference to the best causal explanation. Consider Hacking's most compelling argument for entity realism, which is found in his discussion of microscopy.[24] Hacking argues that the technology of light and electron microscopy has advanced to such a stage that we can now have some confidence that we actually 'see' unobservable theoretical entities through high-powered microscopes. The anti-realist counters that we can never be sure that the observations we make through a

22. Hacking (1983), p. 25
23. Ibid, p. 31
24. Ibid, ch. 7

microscope can warrant the same confidence of veracity as direct, unfiltered observation. In particular, we can never be sure that the images we detect are actually present in the object under micro-inspection, and not merely artefacts of the extremely complex mechanisms of the microscope being used.

Hacking's reply to this skeptical challenge is what he calls "the argument from the grid". When red blood platelets are placed under the scrutiny of an electron microscope, small darkened dots are visible in the blood, which Hacking tells us are called 'dense bodies'. Are these 'dense bodies' mere artefacts of the microscopic apparatus, or are we really 'seeing' an unobservable part of reality? Hacking claims that there is an obvious independent test of the matter. We can fix a numbered microscopic grid on the platelet and place it, in turn, under electron and phosphorescent microscopes, photographing the image in each instance. It turns out that the 'micrographs' depict an identical configuration of 'dense bodies'. Similarly, the numbered grids used are actually photographically reduced copies of large observable grids drawn with pen and ink. When we observe the grid alone, through almost any sort of microscope, the same shapes and letters appear as were originally drawn on a macroscopic scale. We can see how this would tend to make the suspicion that the images of microscopes are mere artefacts rather than genuine features of the world highly implausible. Hacking drives home the point of the example when he writes,

> It would be a preposterous coincidence if time and time again, two completely different physical processes produced identical visual configurations

> which were, however, artefacts of the physical processes rather than real structures in the cell.[25]

And as for the grids themselves,

> To be an anti-realist about the grid, you would have to invoke a malign Cartesian demon of the microscope.[26]

In other words, the best explanation for the same image appearing in all of the different microscopes is just that they are all caused by actual independently existing structural properties of the object being examined--inference to the best causal explanation.

Given this interpretation of the argument from the grid, it is not surprising that van Fraassen objects that it invokes an inferential principle (inference to the best explanation) that he does not accept. Concerning Hacking's claim that the rejection of the argument from the grid would be to invite Cartesian skepticism, van Fraassen writes,

> To add that agnosticism on this point would require a Cartesian demon of the microscope reveals only the unstated premise that persistent similarities in the phenomena *require, must have*, a true explanation. But reliance on that premise is exactly what the previous section [against inference to the best explanation] denied. [my insert][27]

Van Fraassen does not accept such arguments because, as I argued at length in chapter three, he does not think explanation is a ground for belief. I suggest that a resolution of this stand-off will ultimately

25. Hacking (1983), p. 201
26. Ibid, p. 203
27. van Fraassen (1985), p. 298

require an evaluation of Cartwright's claim that causal explanations, unlike nomothetic explanations, have an inherent existential component, and consequently *do* supply grounds for belief. In any event, I think it should be concluded that whatever the value of Cartwright's brand of causal explanationism, Hacking needs it in order for his interventionist arguments to have any force against an anti-realist of van Fraassen's ilk.

It is interesting to notice in this regard that a form of entity realism founded entirely on interventionist arguments will necessarily be far narrower than an entity realism defended on explanationist grounds. For the class of theoretical entities which can be manipulated in experiment forms only a part of the class of theoretical entities which most realists would claim we are justified in believing to exist. For example, black holes, planets orbiting distant stars, and quasars are all--for obvious reasons--not candidates for experimental manipulation. The causal explanationist, on the other hand, can plausibly include such entities in her ontology on the grounds that they are postulated to account for certain phenomena identified by astronomers. And even among the microscopic entities of contemporary physics, not all are presently up for direct manipulation. Free quarks and photons, for example, enjoy extensive empirical support, but cannot yet be put to use in experimental design. In general these sorts of entities are the "more hypothetical parts of nature",[28] which, according to Hacking, we investigate by manipulating comparatively robust entities like

28. Hacking (1983), p. 265

electrons. Again, our belief in these entities, if it is to be justified at all, must be justified by explanationist means in the reasoning from effect to cause.

Towards the end of Representing and Intervening, Hacking makes a concession on this point: "The experimental argument for realism does not say that only the experimenters objects exist."[29] If what I have been arguing is correct, the best argument we can have for believing that the others exist as well must take the form of Cartwright's causal explanationism. And Hacking would do well to reconsider his wholesale dismissal of this sort of realist strategy. I conclude that a throughgoing and integral realism about theoretical entities, like all previous versions of scientific realism, will find its most promising philosophical defense in a version of the principle of inference to the best explanation.

III. Entity Realism and Scientific Progress

In this thesis I have concentrated almost entirely on epistemological issues within the philosophy of science. Specifically, I have tried to show that the realist - anti-realist debate about the extent of scientific knowledge ultimately rests on questions concerning the legitimacy of the principle of inference to the best explanation. I have argued that of the several forms of scientific realism that have been defended by the employment of this principle, only entity realism survives the various objections to explanatory inference that have been brought forth by the anti-

29. Hacking (1983), p. 275

realist camp.

Even within the range of epistemological philosophy of science, the problems I have been dealing with are not the whole story. A much broader epistemological issue concerns the nature of scientific progress. Before the second half of this century, questions about the character of scientific progress were immediately dismissed as non-starters. At that time, science was thought to embody the ultimate method of rational inquiry into the nature of things. As a result, science was typically regarded as the paradigm of intellectual progress against which progress in all other cognitive endeavours was to be measured. But with increased attention on the part of philosophers to the actually history of science, and with the emergence of sophisticated post-positivist anti-realist theories of science from philosophers such as van Fraassen and Laudan, this simple idyllic picture of scientific progress has come under severe and varied criticism. Questions about what the appropriate model of scientific progress would look like, and about whether or not science is indeed progressing towards objective knowledge of the world, now constitute genuine and open questions in the philosophy of science.

One's view of scientific progress will be influenced to a large extent by one's position in the realist - anti-realist debate. On the extreme anti-realist end of the spectrum, sociologists of knowledge such as Barry Barnes argue that scientific progress must be understood in terms of the social structure of the scientific community and the like.[30] Less radically, anti-realist philosophers

30. See for example Barnes (1974)

tend to think of scientific progress in purely empirical terms. Scientific realists on the other hand, construe scientific progress in a more robust epistemological sense, as somehow getting at the truth about the world both at the observable and the unobservable level. The basic idea for the realist here is that science is progressive in a way that other areas like theology, literature, and perhaps philosophy, are not. This realist sentiment is well expressed by Thomas Kuhn in the postscript to The Structure of Scientific Revolutions, a work that was probably the most influential in upsetting the traditional view of scientific progress:

> Though scientific development may resemble that in other other fields more closely than has been supposed, it is also strikingly different. To say, for example, that the sciences, at least after a certain point in their development, progress in a way that other fields do not, cannot have been all wrong, whatever science itself may be.[31]

In this section I want to speculate briefly on the sort of account of scientific progress that must emerge from a view of science that is guided by the epistemological perspective of entity realism. The subject is best approached in relation to a standard objection to entity realism--that our knowledge of theoretical entities is inextricably tied into the basic principles and concepts of one or another theory. In many instances it may be impossible to separate the entities in any sharp way from the basic presuppositions and laws of the theory. In this sense, it is difficult to see how one could

31. Kuhn (1970), p. 209

sustain a strong realist thesis about theoretical entities, while remaining wholly skeptical about the basic claims of the theory in which the entities are conceived. How can one be anti-realist about the theories when it is the theories themselves that determine the character of the theoretical entities postulated?

One answer available to the entity realist can be derived from the recent history of science, which suggests that a good number of theoretical entities persist through theory change. For example, beginning with the research in cathode ray charges undertaken by J. J. Thomson, the electron has appeared in the various different theories of Lorentz, Rutherford, and Bohr. Similarly with atoms, genes, and positrons. Not all entities which function in causal explanations will be radically theory-dependent. Indeed it is precisely those entities whose existence is experimentally confirmed through various theoretical contexts that we will have the most confidence in. And while it is true that different theories employing entities like electrons may say rather different things about the causal properties of the entity, it is not the electron of any particular theory in which the entity realist recommends belief. Rather it is the electron--whose exact character is as yet not completely known--but the existence of which we have excellent reason to believe, whatever its exact character. In response to van Fraassen's rhetorical question as to precisely whose electron the entity realist is committed to, Cartwright answers, "it is the electron, about which we have a large number of incomplete and sometimes conflicting theories."[32] The

32. Cartwright (1983), p. 92

point is that theoretical entities tend to survive even through the rejection of the fundamental theoretical laws and equations governing their behaviour. It is therefore no knock-down argument against entity-realism that theoretical entities are usually postulated to serve a particular theory.

I suspect that a final answer to this sort of objection will depend ultimately on the legitimacy of Cartwright's distinction between the theoretical explanations of fundamental laws, and the causal explanations that employ theoretical entities. For it is the position of the entity realist that abstract theories and concrete entities both have important explanatory functions in scientific inquiry. But because only the latter involve causal explanations, while the former are nomothetic, the pursuit of explanation garners knowledge of unobservable entities, but not of the basic laws of nature. Therefore, however the central concepts of electrons and the like are arrived at in theory formulation, they automatically garner a distinct epistemological status from the fundamental laws and equations of the theories themselves. The status of theoretical entities is to be judged by one criterion: their success in isolated causal explanations; the status of theoretical laws by another criterion: their success in logically classifying and summarizing the disparate phenomena. Only success relative to the first criterion warrants epistemological conclusions. It is this strict division between the explanatory character of scientific theories, and that of the entities they postulate, that grounds the half-realist/half anti-realist epistemology of the entity realist. So long as one accepts this basic distinction, it is not inconsistent to believe in the entities of a certain theory, but not the basic equations of the theory in which statements about the behaviour of the entity are couched.

We can derive from this discussion an idea of the kind of account of scientific progress that will be adopted by the entity realist. Roughly, she will contend that science progresses in two quite different respects. The first respect is essentially anti-realist and pragmatic in tone. Science progresses, and has progressed, to the extent that it has been able to produce simple and abstract laws covering an increasingly broad range of 'complicated and messy' phenomena. Progress in this sense is non-realistic. The increasing explanatory power of the basic equations is to be regarded as a pragmatic advance, but not as an increase in knowledge in any traditional sense. The second respect in which the entity realist sees science as progressive is avowedly realist. Science has progressed in the sense that it has provided an increasingly more accurate account of the causal properties of unobservable entities. Through scientific inquiry we have genuinely increased our knowledge of the nature of the unobservable. But we have not thereby achieved, and indeed should not expect to achieve, a true story about the ultimate laws of nature dictating the behaviour of these entities.

This view of scientific progress is radically at odds with both the unity of science principle that was popular among philosophers in the Vienna Circle and afterwards, and with the idea, currently in vogue among some theoretical physicists such as Stephen Hawking, that we are on the verge of identifying the ultimate true laws of everything. The unity of science principle says that eventually we will find basic laws that are true in all of the various sciences of psychology, biology, chemistry, etc. Once this has been accomplished, it will be possible to effect a reduction of the various laws of these special sciences to the fundamental equations of physics. The entity realist shares no such confidence about the unity of science. For she does not even believe that the fundamental laws of physics are, or will be, true of the objects in its own domain, let alone the entities of psychology, biology and chemistry. As for unificationism in physics itself, the entity realist is skeptical. All of the research in theoretical physics over the last fifty or so years has pointed to disunity. We should therefore not hold our breath waiting for true laws that will unify all of the phenomena. Cartwright writes,

> How unified is our knowledge? Look at the catalogue of a science or engineering school. The curriculum is divided into tiny, separate subjects that irk the interdisciplinist. Our knowledge of nature, nature as we best see it, is highly compartmentalized. Why think nature itself is unified?[33]

In the same vein, Hacking remarks, "The ideal end of science is not unity but plethora."[34] The actual development of science, as well as the nature of the explanatory function of basic theoretical laws, suggests that the hope of arriving at the basic true laws of all creation is a vain one.

The entity realist conception of scientific progress shares with traditional realism the view that science is giving us an increasingly

33. Cartwright (1983), p. 13
34. Hacking (1983), p. 218

accurate account of the concrete entities and complicated processes that underlie the observable realm of everyday life. On the other hand it shares with traditional anti-realism a profound skepticism about science ever laying down the ultimate true theory of nature's workings.

CONCLUSION

In the introduction I characterized this thesis as a 'rational reconstruction' of the history of the realist - anti-realist debate in recent philosophy of science. We have seen that from this perspective the exchange of arguments leads finally to entity realism, a view that falls roughly between the original realist and anti-realist positions. To the extent to which my reconstruction of the debate has been truly rational, we are led inexorably to the conclusion that entity realism presently constitutes the most rational epistemological view of science available.

It now seems to me that I might equally have written the thesis in the form of a dialogue. We can imagine two interlocutors setting out to determine the nature and extent of scientific knowledge. One (the realist) is in the grips of a moderate kind of scientism and is convinced that our best scientific theories constitute a literally true description of the underlying structure and behaviour of 'reality'. The other (the anti-realist) harbours strong skeptical tendencies and denies that the realist's convictions can ever be philosophically justified.

In the first stage of the dialogue they attempt to isolate the exact nature of their disagreement, and set to one side the issues about which they agree, so as to avoid as much as possible talking past each other. They decide that their quarrel is of an essentially epistemological variety, and in particular centers on the realist's steadfast adherence to principles of explanatory inference.

In the second stage of the dialogue the anti-realist offers dual logical and historical skeptical challenges to the realist's simple inference form the empirical success of scientific theories to their truth. As a result the realist is forced to retreat to a second-order inference from the overall increasing empirical success of science to the increasing verisimilitude of scientific theories, thus preserving a weakened--though still uniquely realist--thesis about scientific progress.

In the third stage of the dialogue the anti-realist charges that even this attenuated realist position must rely on a view of explanation peculiar to realism--that explanation is an epistemological virtue of theories which provides grounds for belief. And the anti-realist objects that this begs the question since he understands explanation as an irreducibly pragmatic feature of scientific theories--grounds for acceptance, but never for belief.

The deadlock is finally resolved in the fourth and final stage of the dialogue when it is discovered that the pragmatic aspects of explanation do not support agnosticism where causal explanations are concerned. And since the explanans of causal explanations in scientific contexts are typically concrete theoretical entities, the interlocutors finally end in the adoption of the 'compromise' position of entity realism. They conclude their dialogue with some speculations about the consequences of this new epistemology of science for the philosophical theory of scientific progress.

As is probably obvious, the four stages of this imaginary dialogue correspond to the four chapters of my thesis. The suggestion that entity realism can be understood as a 'compromise' between the realist and anti-realist perhaps sounds a bit un-philosophical. Perhaps we might better cast the dialogue in terms of Hegelian dialectical logic. From this perspective, the doctrine of entity realism may be viewed as the final rational synthesis of the thesis of realism and its antithesis, anti-realism. Rather than 'compromise' the process is one of sublation--entity realism preserves what is rational about the original realist position, and cancels those aspects of it that anti-realism has identified as irrational.

BIBLIOGRAPHY

Achinstein, Peter, (1968), Concepts of Science, John Hopkins Press.

________________, (1971), Law and Explanation, Clarendon Press.

Barnes, Barry, (1974), Scientific Knowledge and Sociological Theory, Routledge & Kegan Paul.

Bhaskar, Roy, (1975), A Realist Theory of Science, Leeds Books Ltd.

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

_____________, (1984), "The Current Status of Scientific Realism", in Leplin, (1984).

_____________, (1985), "Les Orandi est Les Credendi", in Churchland and Hooker, (1985).

Brown, James R., (1985), "Explaining the Success of Science", Ratio 27, pp. 49-66.

Cartwright, Nancy, (1983), How the Laws of Physics Lie, Clarendon Press.

Churchland, Paul M. and Hooker, Clifford A., (eds.), (1985), Images of Science, The University of Chicago Press.

Clark, Romane, (1982), "When is a Fallacy Valid? Reflections on Backward Reasoning", Notre Dame Journal of Formal Logic 23, pp. 1-13.

Collins, Arthur W., (1966), "Explanation and Causality", Mind 75, pp. 482-500.

Creary, Lewis G., (1981), "Causal Explanations and the Reality of Natural Component Forces", Pacific Philosophical Quarterly 62, pp. 148-157.

Devitt, Michael, (1984), Realism and Truth, Princeton University Press.

______________, (1987), Does Realism Explain the Success of Science?", Review of International Philosophy 41, pp. 29-44.

Duhem, Pierre, (1954), The Aim and Structure of Physical Theory, Princeton University Press.

_____________, (1969), To Save the Phenomena, University of Chicago Press.

Ellis, Brian, (1979), Rational Belief Systems, Oxford University Press.

__________, (1985), "What Science Aims To Do", in Churchland and Hooker, (1985).

Feyerabend, Paul K., (1981a), Realism, Rationalism, and Scientific Method, Cambridge University Press.

________________, (1981b), Problems of Empiricism, Cambridge University Press.

Fine, Arthur, (1984a), "And Not Anti-Realism Either", Nous 18, pp. 51-66.

____________, (1984b), "The Natural Ontological Attitude", in Leplin, (1984).

____________, (1986), "Unnatural Attitudes: Realist and Instrumentalist Attachments to Science", Mind 95, pp. 149-147.

Garfinkel, Alan, (1981), Forms of Explanation, Yale University Press.

Glymour, Clark, (1971), "Theoretical Realism and Theoretical Equivalence", Boston Studies in the Philosophy of Science 8, pp. 275-288.

________________, (1980), Theory and Evidence, Princeton University Press.

Grunbaum, Adolph, (1960), "The Duhemian Argument", Philosophy of Science 27, pp. 75-87.

Gutting, Garry, (1985), "Realism Versus Constructive Empiricism: A Dialogue", in Churchland and Hooker, (1985).

Hacking, Ian, (ed.), (1981), Scientific Revolutions, Oxford University Press.

___________, (1983), Representing and Intervening, Cambridge University Press.

Hanson, Norwood Russell, (1958), Patterns of Discovery, Cambridge University Press.

Hardin, C. L. and Rosenberg, A., (1982), "In Defense of Convergent Realism", Philosophy of Science 49, pp. 604-615.

Harman, Gilbert, (1965), "Inference to the Best Explanation", Philosophical Review 74, pp. 88-95.

Harre, Rom, (1986), Varieties of Realism, Basil Blackwell.

Hempel, Carl G. and Oppenheim, P., (1948), "Studies in the Logic of Explanation", Philosophy of Science 15, pp. 135-175.

Hempel, Carl G., (1965), Aspects of Scientific Explanation, Macmillan Publishing.

Hesse, Marry., (1974), Structure of Scientific Inference, Macmillan Press.

Horwich, Paul, (1982), "How to Choose Between Empirically Indistinguishable Theories", Journal of Philosophy 79, pp. 61-77.

Jensen, U. J. and Harre, Rom, (eds.), (1981), The Philosophy of Evolution, Harvestor Press.

Kitcher, Philip, (1978), "Theories, Theorists and Theoretical Change", The Philosophical Review 87, pp. 519-547.

Kuhn, Thomas, (1970), Structure of Scientific Revolutions, University of Chicago Press.

Lakatos, Imre, (1970), "History of Science and Its Rational Reconstructions", Boston Studies in the Philosophy of Science, VIII, pp. 91-137.

Laudan, Larry, (1977), Progress and Its Problems, University of California Press.

_____________, (1984a), "A Confutation of Convergent Realism", in Leplin, (1984).

_____________, (1984b), "Realism Without the Real", Philosophy of Science 51, pp. 156-162.

Leplin, Jarret, (ed.), (1984), Scientific Realism, University of California Press.

Mackinnon, E. A., (ed.), (1972), Problems of Scientific Realism, Apple-Century-Crofts.

McGlaughlin, Robert, (ed.), (1982), What?, Where?, When?, Why?, D. Reidel Publishing Co.

McMichael, A., (1985), "van Fraassen's Instrumentalism", British Journal for the Philosophy of Science 36, pp. 257-272.

McMullin, Ernan, (1978), "Structural Explanation", American Philosophical Quarterly 15, pp. 139-147.

________________, (1984), "A Case For Scientifc Realism", in Leplin, (1984).

________________, (1985), "Truth and Explanatory Success", Proceedings of The American Catholic Philosophical Association 59, pp. 206-231.

Musgrave, Alan, (1985), "Realism Versus Constructive Empiricism", in CHurchland and Hooker, (1985).

Newton-Smith, William H., (1978), The Underdetermination of Theory by Data", Aristotelian Society Supplementary Volume 52, pp. 71-91.

______________________, (1981a), "In Defense of Truth", in Jensen and Harre, (1981).

______________________, (1981b), The Rationality of Science, Routledge and Kegan Paul.

Peacocke, Arthur R., (1984), Intimations of Reality, University of Notre Dame Press.

Peirce, Charles S., (1934), Collected Papers, Harvard University Press.

Poincare, Henri, (1958), The Value of Science, Dover Publications Inc.

Putnam, Hilary, (1975), Mathematics, Matter and Method, Cambridge University Press.

Quine, Willard V., (1975), "On Empirically Equivalent Systems of the World", Erkenntnis 9, pp. 313-328.

Reichenbach, Hans, (1938), Experience and Prediction, University of Chicago Press.

Rescher, Nicholas R., (1978), Peirce's Philosophy of Science, University of Notre Dame Press.

Salmon, Wesley C., (1984), Scientific Explanation and the Causal Structure of the World, Princeton University Press.

Schlesinger, George, (1974), Confirmation and Confirmability, Oxford University Press.

Sellars, Wilfred, (1962), Science, Perception and Reality, Humanities Press.

______________, (1965), "Scientific Realism or Irenic Instrumentalism", Boston Studies in the Philosophy of Science, Volume II, pp. 171-204.

______________, (1976), "Is Scientific Realism Tenable?", Proceedings of the Philosophy of Science Association, 1976, Volume 2, pp. 307-334.

van Fraassen, Bas C., (1977), "The Pragmatics of Explanation", American Philosophical Quarterly 14, pp. 143-150.

____________________, (1980), The Scientific Image, Clarendon Press.

____________________, (1985), "Empiricism in the Philosophy of Science", in Churchland and Hooker, (1985).

Woodward, James, (1986), "Are Singular Causal Explanations Implicit Covering - Law Explanations?", Canadian Journal of Philosophy 16, pp. 253-280.

Worral, John, (1982), "Scientific Realism and Scientific Change", Philosophical Quarterly 32, pp. 201-231.

Wylie, Alison, (1986), Arguments for Scientific Realism: The Ascending Spiral", American Philosophical Quarterly 23, pp. 287-298.